



Lawry, J. (2014). Probability, fuzziness and borderline cases.  
*International Journal of Approximate Reasoning*, 55(5), 1164–1184.  
<https://doi.org/10.1016/j.ijar.2014.01.004>

Peer reviewed version

Link to published version (if available):  
[10.1016/j.ijar.2014.01.004](https://doi.org/10.1016/j.ijar.2014.01.004)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# Probability, Fuzziness and Borderline Cases

Jonathan Lawry,  
Department of Engineering Mathematics,  
University of Bristol,  
Bristol BS8 1UB, UK  
email: j.lawry@bris.ac.uk

## Abstract

An integrated approach to truth-gaps and epistemic uncertainty is described, based on probability distributions defined over a set of three-valued truth models. This combines the explicit representation of borderline cases with both semantic and stochastic uncertainty, in order to define measures of subjective belief in vague propositions. Within this framework we investigate bridges between probability theory and fuzziness in a propositional logic setting. In particular, when the underlying truth model is from Kleene’s three-valued logic then we provide a complete characterisation of compositional min-max fuzzy truth degrees. For classical and supervaluationist truth models we find partial bridges, with min and max combination rules only recoverable on a fragment of the language. Across all of these different types of truth valuations, min-max operators are resultant in those case in which there is *only* uncertainty about the relative sharpness or vagueness of the interpretation of the language.

**Keywords** Vagueness, Truth-gaps, Probability, Fuzziness.

## 1 Introduction

From its inception fuzzy logic has largely been seen as a direct competitor to probability as a theory of uncertainty. This perceived conflict has motivated several vigorous critiques of fuzzy theory by Bayesian statisticians including Lindley [35], Cheeseman [2], [3] and, Laviolette and Seaman [26], amongst others. As particularly emphasised by Cooke [6], a central theme of many of these critiques is that of *operational semantics*. Consider, for example, a witness to a robbery describing the suspect to a police officer. If she says that ‘the suspect is short’ with fuzzy truth degree 0.6, then what does she mean? What exactly can the police officer infer from the value 0.6? The benchmark in this context is generally taken to be the work of de Finetti [7] who, inspired by the operationalism movement in physics, proposed betting behaviour as an operational semantics for subjective probability (See Paris [40] for an exposition). It is certainly the case that de Finetti’s results lend weight to the claim that probability is the optimal calculus for measures of uncertainty,

since agents adopting any other type of measure would be subject to Dutch books <sup>1</sup>. In other words, to quote Lindley [35], we would seem to be forced to concede that ‘the only satisfactory description of uncertainty is probability’.

Yet this argument is not quite as compelling as it might at first seem. Although widely accepted, de Finetti’s betting model does make certain assumptions which are open to question. For example, in the formulation it is assumed that the buying and selling prices of a bet are the same. However, Walley [52] introduces a more general framework of coherent gambles which does not make such an assumption and naturally results in lower and upper probabilities as measures of subjective belief. Furthermore, Paris [41] observes that bets are inherently defined relative to an underlying truth-model. In the case of a Tarskian (or classical) truth-model then to avoid Dutch books the resulting measure of subjective belief must be a probability measure on the set of sentences of the relevant language i.e. on those sentences about which bets are offered. In addition, this turns out to be equivalent to defining a probability distribution on the set of possible truth-valuations of the language, and then taking the belief value of a sentence to be the probability of those valuations for which it is true. Indeed, Paris [41] has shown that if, more generally, the outcome of bets are dependent on a truth-valuation selected from a set of binary functions  $\mathbb{B}$ , then the avoidance of Dutch bets requires that belief measures be similarly defined in terms of probability distributions over  $\mathbb{B}$ . However, depending on the nature of these binary functions  $\mathbb{B}$ , then the resulting measure may not be a probability measure over the sentences of the language. This latter point is important since it will allow us to give a probabilistic interpretation to the fully truth-functional min-max calculus of fuzzy logic, with an associated betting semantics, but where probabilities are defined over a set of non-Tarskian truth-models.

We claim that non-Tarskian truth-valuations are particularly relevant when we allow vague propositions into our language. To make this case we must first identify some important characteristics of vagueness in natural language. Both Lindley [35] and Cheeseman [3] argue that quantifying subjective belief in vague propositions such as ‘the suspect is short’ presents no significant difficulty to probability theory. More specifically, Cheeseman [3] proposes that vagueness should be understood as uncertainty concerning the underlying interpretation of the language, e.g. uncertainty about what is the exact definition of *short*, and consequently is probabilistic in nature. The idea that uncertainty about meaning should be explicitly quantified is clearly an important and rather neglected one. To a great extent language is learned empirically through interactions with others and the environment, as well as through exposure to social media such as the written word. Consequently, the association of words and meanings, is not only naturally dynamic but also an emergent phenomenon. There is no central authority or oracle to interrogate about the

---

<sup>1</sup>willingly accepting a sequence of bets which are sure to result in an overall loss, irrespective of the true state of the world

correct boundaries of vague categories. Instead, individual agents only have data about how others have used these categories in similar contexts. Subjective uncertainty must surely then be inherent to any such complex system.

Lawry and Tang [31] refer to uncertainty about meaning as *semantic uncertainty* and contrast it with *stochastic uncertainty* which refers to uncertainty about the state of the world. For example, in order for our robbery witness to assess her level of belief in the proposition ‘the suspect is short’, she would need to take into account both her stochastic uncertainty about the height of the suspect, and her semantic uncertainty about the location of the height boundary that divides *short* from *not short*. We argue in Lawry [27] that semantic uncertainty is effectively epistemic in nature and can be treated probabilistically. This viewpoint is broadly in keeping with that of Lassiter [25] and consistent with both the random set theory and likelihood interpretations of fuzzy membership functions as described in [11]. Now if we accept that the underlying truth-model should be Tarskian even for vague propositions, then this epistemic view of semantic uncertainty brings us close to Williamson’s epistemic theory of vagueness [54]. In this paper, however, we will argue that there is more to vagueness than *only* semantic uncertainty and therefore that a more flexible truth theory may be appropriate when evaluating vague propositions.

Vagueness is a multifaceted phenomenon. For instance, Keefe and Smith [22] identify ‘three interrelated features’ of vague predicates; 1) borderline cases 2) lack of well-defined extensions or boundaries, and 3) susceptibility to sorites paradoxes. Arguably 2) relates to semantic uncertainty as outlined above and is amenable to probabilistic treatment. The sorites paradoxes in 3) concern sequences of indistinguishable elements where the first element in the sequence belongs to a certain class and the last to its negation. A classical model then requires drawing an arbitrary boundary between the class and its negation which by necessity then allocates indistinguishable elements to opposite classes. We will not address this aspect of vagueness in this current paper except to the extent that is related to and influenced by features 1) and 2), and we would point interested readers towards the extensive literature on the topic, an excellent overview of which is given in Williamson [54]. Instead, we aim to describe and justify a model of subjective belief for vague propositions in which feature 1) plays a central role. More specifically, we will assume a three-valued truth model in which borderline cases naturally occur as a result of vague predicates being only partially (or incompletely) defined.

In the sequel we propose an integrated approach to truth-gaps and epistemic uncertainty based on probability distributions defined over a finite set of three-valued truth models for a propositional language. This will combine the explicit representation of borderline cases with a quantification of both semantic and stochastic uncertainty, so as to define measures of subjective belief over the sentences of a language containing vague propositions. As well as suggesting a general framework for the treatment of both vagueness and uncertainty, this approach will provide an number of bridges, in the sense of

[10], between probability theory and fuzzy logic, including a complete characterisation of truth-functional min-max fuzzy logic [55]. Given the broad scope of both subjects our bridges are by necessity narrow ones. For example, we have as yet made no attempt to investigate the wide variety of truth-functional calculi developed by fuzzy logicians [19]. Instead, the proposed bridges will only be with Zadeh’s original min-max calculus [55]. Despite this limitation we would argue that the proposed probabilistic semantics for fuzzy logics goes some way to answering the interpretational questions posed in [2], [3] [6], [26] and [35] .

An outline of the paper is as follows: Section 2 introduces valuation pairs as a generic three-valued truth model and then focuses on Kleene’s three-valued logic as a particular case. Section 3 describes the lower and upper measures of subjective belief which naturally result from defining probability distributions over Kleene valuation pairs. In section 4 we propose a notion of truth degree as the mid-point of the lower and upper measures defined in section 3. Furthermore, we show that truth degrees defined in this way, characterise min-max fuzzy truth degrees in the case when uncertainty only concerns the level of vagueness of the underlying interpretation of the language. In section 5 we outline a betting semantics for truth degrees exploiting Paris’ [41] generalization of de Finetti’s famous result. We then extend our proposed framework to other truth models and, in particular, in section 6 we investigate bridges between fuzzy logic and probabilities defined over both Tarski (classical) and supervaluationist truth models. Finally, section 7 presents some discussion and conclusions. Overall, the work presented builds on and exploits the general framework for subject belief in the presence of vagueness described in [29], [30], [31], [32] and [33].

## 2 Truth-Gaps in a Propositional Language

In the literature two main approaches to modelling truth-gaps have been proposed; three-valued logic [23] and supervaluationism [16]. Here we will introduce a generic form of truth valuation for propositional languages which is sufficiently flexible to allow us to represent both these types of models in their simplest form. This will enable us to study both approaches using the same basic notation, thus making it easier to identify and discuss relationships and differences between them. We also claim that the proposed notion of valuation pair is an intuitive way to represent truth-gaps, which when combined with probability, naturally leads to lower and upper subjective belief measures on the sentences of the language.

Let  $\mathcal{L}$  be a language of propositional logic with connectives  $\wedge$ ,  $\vee$  and  $\neg$  and propositional variables  $\mathcal{P} = \{p_1, \dots, p_n\}$ . Let  $S\mathcal{L}$  denote the sentences of  $\mathcal{L}$  as generated recursively from the propositional variables by application of the three connectives, and let  $L\mathcal{L} = \mathcal{P} \cup \{\neg p : p \in \mathcal{P}\}$  denote the literals of  $\mathcal{L}$ . A valuation pair on  $S\mathcal{L}$  then consists of two binary functions  $\underline{v}$  and  $\bar{v}$  representing lower and upper truth values. The underly-

ing idea is that  $\underline{v}$  represents the strong criterion of *absolutely true* while  $\bar{v}$  represents the weaker criteria of *not absolutely false*.

**Definition 1.** *Valuation Pairs*

A valuation pair is a pair of functions  $\vec{v} = (\underline{v}, \bar{v})$  where  $\underline{v} : S\mathcal{L} \rightarrow \{0, 1\}$  and  $\bar{v} : S\mathcal{L} \rightarrow \{0, 1\}$  such that  $\underline{v} \leq \bar{v}$ . Furthermore,  $\forall \theta, \varphi \in S\mathcal{L}$ , if  $\underline{v}(\theta) = \bar{v}(\theta) = \alpha$  and  $\underline{v}(\varphi) = \bar{v}(\varphi) = \beta$  then  $\underline{v}(\neg\theta) = \bar{v}(\neg\theta) = 1 - \alpha$ ,  $\underline{v}(\theta \wedge \varphi) = \bar{v}(\theta \wedge \varphi) = \min(\alpha, \beta)$  and  $\underline{v}(\theta \vee \varphi) = \bar{v}(\theta \vee \varphi) = \max(\alpha, \beta)$ .

Notice that we can also think of a valuation pair as a three-valued mapping with  $\vec{v}(\theta)$  having possible values  $\mathbf{t} = (1, 1)$ ,  $\mathbf{b} = (0, 1)$  and  $\mathbf{f} = (0, 0)$ , standing for *absolutely true*, *borderline* and *absolutely false* respectively. We will use this three-valued notation interchangeably with the lower and upper valuation notation throughout this paper. The second part of definition 1 is motivated by the intuition that if we restrict ourselves to the non-vague (crisp) sentences of  $\mathcal{L}$  then valuation pairs should have the same properties as classical valuations. Indeed, classical valuations can be viewed as a special case of valuation pairs where for every sentence  $\theta \in S\mathcal{L}$ ,  $\underline{v}(\theta) = \bar{v}(\theta)$ , i.e. as valuations for which the truth value  $\mathbf{b}$  does not occur.

In accordance with [39], we might think of a sentence being absolutely true as meaning that it can be uncontroversially asserted without any risk of censure, while being not absolutely false only means that it is acceptable to assert i.e. one can get away with such an assertion. Recall, our example of a witness describing the suspect as being *short*. If she was to testify to this in a court of law then depending on the actual height of the suspect her statement may be deemed as clearly true or clearly false, in which latter case the witness could be accused of perjury. However, there will also be an intermediate height range for which, while there may be doubt and differing opinions concerning the use of the description *short*, it would not be deemed as definitely inappropriate and hence the witness would not be viewed as committing perjury. In other words, for certain height values of the suspect, it may be acceptable to assert the statement  $p = \text{'the suspect is short'}$ , even though this statement would not be viewed as being absolutely true. One possible model of the predicate *short* exhibiting such truth-gaps could be as follows: Let  $h$  be the height of the suspect and suppose that *short* is defined in terms of lower and upper thresholds  $\underline{h} \leq \bar{h}$  on heights. We might also think of this as a partial interpretation of the predicate *short* generated as an abstraction from independent consideration of certain clearly positive and clearly negative examples. In this case  $p$  is *absolutely true* if  $h \leq \underline{h}$ , *absolutely false* if  $h > \bar{h}$  and *borderline* if  $\underline{h} < h \leq \bar{h}$  (see figure 1).

It is important to note that in this model truth-gaps corresponding to different lower and upper truth valuations are not the result of epistemic uncertainty concerning the state of the world, but are rather due to inherent flexibility in the underlying language conventions. In other words, a truth-gap (or middle truth value in three-valued logic) does

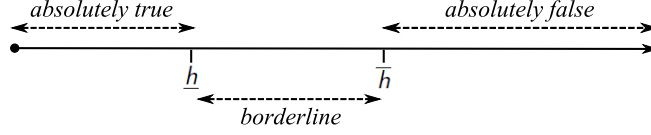


Figure 1: An interpretation of the predicate short incorporating a truth-gap.

not represent an *uncertain* epistemic state [5]. For example, given absolute certainty about the suspect's height  $h$ , the proposition  $p$  may then be *known* to be *borderline* because of the inherent flexibility (or vagueness) in the definition of the concept short i.e. because  $\underline{h} < h \leq \bar{h}$ . The potential confusion resulting from applying many-valued logic to model epistemic uncertainty is highlighted by Dubois in [12]. Here we emphasize the truth-value status of the intermediate case by the use of the term *borderline* rather than ‘uncertain’ or ‘unknown’ as originally suggested by Kleene [23].

Initially, we will focus on valuation pairs which capture Kleene's three-valued logic [23], and then return in section 6 to investigate supervaluations within this framework.

**Definition 2.** *Kleene Valuation Pairs*

A Kleene valuation pair on  $\mathcal{L}$  is a valuation pair  $\vec{v} = (\underline{v}, \bar{v})$  where  $\forall \theta, \varphi \in S\mathcal{L}$  the following hold:

- $\underline{v}(\neg\theta) = 1 - \bar{v}(\theta)$  and  $\bar{v}(\neg\theta) = 1 - \underline{v}(\theta)$
- $\underline{v}(\theta \wedge \varphi) = \min(\underline{v}(\theta), \underline{v}(\varphi))$  and  $\bar{v}(\theta \wedge \varphi) = \min(\bar{v}(\theta), \bar{v}(\varphi))$
- $\underline{v}(\theta \vee \varphi) = \max(\underline{v}(\theta), \underline{v}(\varphi))$  and  $\bar{v}(\theta \vee \varphi) = \max(\bar{v}(\theta), \bar{v}(\varphi))$

Let  $\mathbb{V}_k$  denote the set of all Kleene valuation pairs on  $\mathcal{L}$ .

From definition 2 we can generate truth tables for the connectives  $\wedge$ ,  $\vee$  and  $\neg$  in terms of the truth values  $\{\mathbf{t}, \mathbf{b}, \mathbf{f}\}$  identical to those of Kleene's logic [23] (see table 1)<sup>2</sup>.

|          |          |  |          |          |          |          |  |          |          |          |          |
|----------|----------|--|----------|----------|----------|----------|--|----------|----------|----------|----------|
|          | $\neg$   |  | $\wedge$ | <b>t</b> | <b>b</b> | <b>f</b> |  | $\vee$   | <b>t</b> | <b>b</b> | <b>f</b> |
| <b>t</b> | <b>f</b> |  | <b>t</b> | <b>t</b> | <b>b</b> | <b>f</b> |  | <b>t</b> | <b>t</b> | <b>t</b> |          |
| <b>b</b> | <b>b</b> |  | <b>b</b> | <b>b</b> | <b>b</b> | <b>f</b> |  | <b>b</b> | <b>t</b> | <b>b</b> | <b>b</b> |
| <b>f</b> | <b>t</b> |  | <b>f</b> | <b>f</b> | <b>f</b> | <b>f</b> |  | <b>f</b> | <b>t</b> | <b>b</b> | <b>f</b> |

Table 1: Truth tables from Kleene's strong three-valued logic.

Kleene's logic has been proposed as an appropriate formalism for truth gaps by a number of authors including Shapiro [45], Soames [47] and Tye [49]. In some respects it is a rather conservative calculus which tends to allocate *borderline* status to sentences more readily than, say, a supervaluation approach. In particular, it permits *borderline*

<sup>2</sup>Recall that in valuation pair notation  $\mathbf{t} = (1, 1)$ ,  $\mathbf{b} = (0, 1)$  and  $\mathbf{f} = (0, 0)$ .

contradictions (and tautologies), since from table 1 we can see that  $\vec{v}(\theta \wedge \neg\theta) = \mathbf{b}$  ( $\vec{v}(\theta \vee \neg\theta) = \mathbf{b}$ ) exactly when  $\vec{v}(\theta) = \mathbf{b}$ . Such a feature is contentious, with some arguing that a contradiction  $\theta \wedge \neg\theta$  should always be deemed absolutely false, irrespective of the truth-value of  $\theta$ . For instance, as we shall see in section 6, supervaluations preserve all classical contradictions, tautologies and equivalences. However, since we are modelling truth values rather than epistemic states [5], the validity or otherwise of borderline contradictions is rather unclear. This is because truth models simply represent conventions governing how the underlying language is interpreted, and as such they have the status of primitives. In particular, for the propositional language  $\mathcal{L}$  the properties we choose for valuation pairs effectively codify our understanding of the connectives  $\wedge$ ,  $\vee$  and  $\neg$ , in a three valued context. We could of course ask which type of valuation provides the best model of vagueness in natural language. This question remains firmly open, due to the fact that there have been very few empirical studies directly addressing issues such as borderline contradictions. One such study, described by Ripley [43], involves an experiment in which the participants are asked to state their level of agreement with contradictions of the form ‘the circle is near the square and it isn’t near the square’, when presented with a sequence of images, in which a circle and square are shown at various distances from each other along a line. Whilst the results are somewhat inconclusive Ripley does find evidence to suggest an increased willingness amongst the participants to accept contradictions in borderline cases.

## 2.1 A Positive and Negative Characterisation of Kleene Valuations

In this sub-section we consider a characterisation of Kleene valuation pairs in terms of positive and negative propositions as represented by the sets of absolutely true propositional variables and absolutely true negated propositional variables respectively. More formally, a Kleene valuation pair  $\vec{v}$  can be characterised by an *orthopair*<sup>3</sup>  $(P, N) \in 2^{\mathcal{P}} \times 2^{\mathcal{P}}$  where  $P = \{p_i \in \mathcal{P} : \underline{v}(p_i) = 1\}$  and  $N = \{p_i \in \mathcal{P} : \underline{v}(\neg p_i) = 1\}$ . Notice, that from definition 2 it holds immediately that  $P \cap N = \emptyset$ . The next result shows how the value of  $\vec{v}$  across  $S\mathcal{L}$  can be determined directly from its associated orthopairs  $(P, N)$ . The following mapping between sentences of  $\mathcal{L}$  and sets of pairs, forms the basis of our proposed characterisation.

### Definition 3. $\lambda$ -mapping

Let  $\lambda : S\mathcal{L} \rightarrow 2^{2^{\mathcal{P}} \times 2^{\mathcal{P}}}$  be defined recursively as follows:  $\forall \theta, \varphi \in S\mathcal{L}$

- $\lambda(p_i) = \{(F, G) \in 2^{\mathcal{P}} \times 2^{\mathcal{P}} : p_i \in F\}$
- $\lambda(\theta \wedge \varphi) = \lambda(\theta) \cap \lambda(\varphi)$
- $\lambda(\theta \vee \varphi) = \lambda(\theta) \cup \lambda(\varphi)$

---

<sup>3</sup>An orthopair is simply a pair of sets  $(F, G)$  such that  $F \cap G = \emptyset$  [4].



- $\lambda(\neg\theta) = \{(G^c, F^c) : (F, G) \in \lambda(\theta)\}^c$

Notice that the  $\lambda$ -mapping in definition 3 is not restricted solely to orthopairs but also includes pairs of sets of propositional variables with non-empty intersection. As described in [29], such sets characterize a more general class of binary function pairs  $(v_1, v_2)$  which satisfy the duality and min-max combination rules of definition 2 but without the requirement that  $v_1 \leq v_2$ <sup>4</sup>. This class of function pairs clearly includes Kleene valuation pairs as a special case. Consequently, many of the results in [29] carry across to the current context including the following characterization theorem.

**Theorem 4.** [29] *For a Kleene valuation pair  $\vec{v} \in \mathbb{V}_k$  characterised by orthopair  $(P, N)$  it holds that  $\forall \theta \in S\mathcal{L}$ ,  $\underline{v}(\theta) = 1$  if and only if  $(P, N) \in \lambda(\theta)$  and  $\bar{v}(\theta) = 1$  if and only if  $(P, N) \in \lambda(\neg\theta)^c$ .*

**Example 5.** *Let  $p_i, p_j \in \mathcal{P}$  then*

$\lambda(p_i) = \{(F, G) : p_i \in F\}$ ,  $\lambda(\neg p_j) = \{(F, G) : p_j \in G\}$  and  $\lambda(p_i \wedge \neg p_j) = \{(F, G) : p_i \in F, p_j \in G\}$ . Hence,  $\underline{v}(p_i) = 1$  iff  $p_i \in P$  and  $\bar{v}(p_i) = 1$  iff  $p_i \notin N$ . Similarly,  $\underline{v}(\neg p_j) = 1$  iff  $p_j \in N$  and  $\bar{v}(\neg p_j) = 1$  iff  $p_j \notin P$ . Furthermore,  $\underline{v}(p_i \wedge \neg p_j) = 1$  iff  $p_i \in P$  and  $p_j \in N$ , and  $\bar{v}(p_i \wedge \neg p_j) = 1$  iff  $p_i \notin N$  and  $p_j \notin P$ .

## 2.2 Semantic Precision: A Vagueness Ordering on Valuation Pairs

We now define *semantic precision* as a natural partial ordering on valuation pairs. This concerns the situation in which one valuation pairs admits more borderline cases than another but where otherwise their truth values agree. More formally, valuation pair  $\vec{v}_1$  is less semantically precise than  $\vec{v}_2$ , denoted  $\vec{v}_1 \preceq \vec{v}_2$ , if they disagree only for some set of sentences of  $\mathcal{L}$ , which being identified as either absolutely true or absolutely false by  $\vec{v}_2$ , are classified as being borderline cases by  $\vec{v}_1$ . In other words,  $\vec{v}_1$  is less semantically precise than  $\vec{v}_2$  if all of the absolutely true and absolutely false valuations of  $\vec{v}_1$  are preserved by  $\vec{v}_2$ . Hence, one might think of  $\preceq$  as ordering valuation pairs according to their relative vagueness. Shapiro [45] proposed essentially the same ordering of interpretations which he refers to as *sharpening* i.e.  $\vec{v}_1 \preceq \vec{v}_2$  means that  $\vec{v}_2$  extends or sharpens  $\vec{v}_1$ .

**Definition 6.** *Semantic Precision*

$\vec{v}_1 \preceq \vec{v}_2$  if and only if  $\forall \theta \in S\mathcal{L}$ ,  $\underline{v}_1(\theta) \leq \underline{v}_2(\theta)$  and  $\bar{v}_1(\theta) \geq \bar{v}_2(\theta)$ <sup>5</sup>.

**Theorem 7.** [29]  $\forall \vec{v}_1, \vec{v}_2 \in \mathbb{V}_k$ , where  $\vec{v}_1$  and  $\vec{v}_2$  are characterised by orthopairs  $(P_1, N_1)$  and  $(P_2, N_2)$  respectively, it holds that;  $\vec{v}_1 \preceq \vec{v}_2$  if and only if  $P_1 \subseteq P_2$  and  $N_1 \subseteq N_2$ .

<sup>4</sup>We are referring here to the more general class of pairs of functions  $(v_1, v_2)$  where  $v_1 : S\mathcal{L} \rightarrow \{0, 1\}$ ,  $v_2 : S\mathcal{L} \rightarrow \{0, 1\}$  and  $\forall \theta, \varphi \in S\mathcal{L}$ ,  $v_1(\neg\theta) = 1 - v_2(\theta)$ ,  $v_2(\neg\theta) = 1 - v_1(\theta)$ ,  $v_1(\theta \wedge \varphi) = \min(v_1(\theta), v_1(\varphi))$ ,  $v_2(\theta \wedge \varphi) = \min(v_2(\theta), v_2(\varphi))$ ,  $v_1(\theta \vee \varphi) = \max(v_1(\theta), v_1(\varphi))$  and  $v_2(\theta \vee \varphi) = \max(v_2(\theta), v_2(\varphi))$ , but where it is not required that  $v_1 \leq v_2$ . In this case  $(v_1, v_2)$  can take four values corresponding to  $\mathbf{t} = (1, 1)$ ,  $\mathbf{b} = (0, 1)$ ,  $\mathbf{i} = (1, 0)$  and  $\mathbf{f} = (0, 0)$ . The truth tables for these four truth values are then those of Belnap's four-valued logic [1].

<sup>5</sup>Alternatively, using three-valued notation  $\vec{v}_1 \preceq \vec{v}_2$  if and only if  $\forall \theta \in S\mathcal{L}$ ,  $\vec{v}_1(\theta) = \mathbf{t} \Rightarrow \vec{v}_2(\theta) = \mathbf{t}$

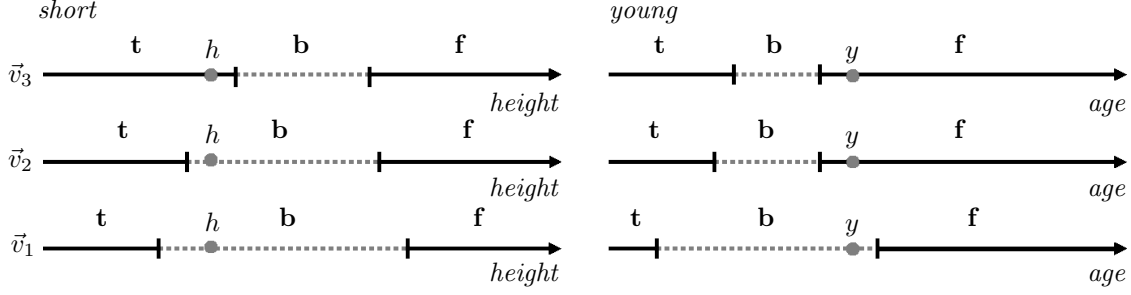


Figure 2: A sequence of valuation pairs  $\vec{v}_1 \preceq \vec{v}_2 \preceq \vec{v}_3$  as characterised by increasingly vague interpretations of the predicates *short* and *young*.

Figure 2 shows three valuation pairs relevant to the crime reporting example described in section 1, each based on different interpretations of the predicates *short* and *young*. Here we now have two propositions corresponding to  $p_1$  = ‘the suspect is short’ and  $p_2$  = ‘the suspect is young’. The predicates *short* and *young* are defined by lower and upper thresholds on height and age respectively, thus identifying clear **t**, **b** and **f** intervals for each predicate, as indicated in figure 2. The height  $h$  and age  $y$  of the suspect are then shown as being constant across the three valuations. Working down the three interpretations the definitions of both predicates become increasingly vague in the sense that the size of each borderline interval is increasing. In orthopair notation the three valuation pairs are  $\vec{v}_1 = (\emptyset, \emptyset)$ ,  $\vec{v}_2 = (\emptyset, \{p_2\})$  and  $\vec{v}_3 = (\{p_1\}, \{p_2\})$  respectively. As such they satisfy the nestedness condition of theorem 7 and hence,  $\vec{v}_1 \preceq \vec{v}_2 \preceq \vec{v}_3$ .

### 3 Probability in the Context of a Three-Valued Truth Model

Within the proposed three-valued framework, we take the view that uncertainty concerning the sentences of  $\mathcal{L}$  effectively corresponds to uncertainty as to which is the correct valuation pair for  $\mathcal{L}$ . A betting argument in favour of this assumption is explored in section 5. As discussed in the introduction, we propose a probabilistic treatment of both semantic and stochastic uncertainty in the context of an underlying three-valued truth model. For example, consider the three-valued interpretation of the predicate *short* as shown in figure 1. In this case semantic uncertainty manifests itself in terms of uncertainty about the exact values of the thresholds  $\underline{h}$  and  $\bar{h}$ , whilst stochastic uncertainty would be about the suspect’s height  $h$ . Treating both types of uncertainty as being epistemic in nature, and defining a joint probability distribution over  $\underline{h}$ ,  $\bar{h}$  and  $h$ , together with other similar variables arising in relation to other vague predicates, would then naturally result in a probability distribution over the valuation pairs of  $\mathcal{L}$ . This integrated treatment of both semantic and stochastic uncertainty is also consistent with the ideas of Lassiter [25] who proposes a joint probability distribution over states of the world and language interpretations, and then outlines a treatment of Sorites based on this approach.

Viewing semantic uncertainty as being epistemic in nature requires that agents make the assumption that there is a correct underlying interpretation of the language  $\mathcal{L}$ , but about which they may be uncertain. This is a weaker version of the epistemic theory of vagueness as expounded by Timothy Williamson [54] and which we refer to as the *epistemic stance* [27]. Williamson’s theory assumes that for a vague predicate there is a precise but unknown boundary between it and its negation. In contrast the epistemic stance corresponds to the more pragmatic view that individuals, when faced with decision problems about what to assert, find it useful as part of a decision making strategy to simply *assume* that there is an underlying correct interpretation of  $\mathcal{L}$ . In other words, when deciding what to assert agents behave as if the epistemic theory is correct. Another difference between the epistemic theory and our current approach is that the former assumes that the underlying truth model is classical, while here we assume a three-valued model which can exhibit truth-gaps.

In the following definition we assume that uncertainty is quantified by a probability measure  $w$  on the set of Kleene valuation pairs  $\mathbb{V}_k$  of  $\mathcal{L}$ .

**Definition 8.** *Kleene Belief Pairs* [29]

Let  $w$  be a probability distribution defined on  $\mathbb{V}_k$  so that  $w(\vec{v})$  is the agent’s subjective belief that  $\vec{v}$  is the true valuation pair for  $\mathcal{L}$ . Then  $\vec{\mu} = (\underline{\mu}, \bar{\mu})$  is a Kleene belief pair where  $\forall \theta \in S\mathcal{L}$ ,

$$\underline{\mu}(\theta) = w(\{\vec{v} : \underline{v}(\theta) = 1\}) \text{ and } \bar{\mu}(\theta) = w(\{\vec{v} : \bar{v}(\theta) = 1\})$$

Alternatively,

$$\underline{\mu}(\theta) = w(\{\vec{v} : \vec{v}(\theta) = \mathbf{t}\}) \text{ and } \bar{\mu}(\theta) = w(\{\vec{v} : \vec{v}(\theta) \neq \mathbf{f}\})$$

Note that here and in the sequel we abuse notation slightly and use the same symbol  $w$  to stand for both a probability distribution on  $\mathbb{V}_k$  and the probability measure on  $2^{\mathbb{V}_k}$  which it then induces.

In this formulation it also trivially holds that:

$$w(\{\vec{v} : \vec{v}(\theta) = \mathbf{f}\}) = \underline{\mu}(\neg\theta) \text{ and } w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\}) = \bar{\mu}(\theta) - \underline{\mu}(\theta)$$

The following theorem highlights a number of properties of Kleene belief pairs, including additivity. This latter property in particular, distinguishes Kleene belief pairs from Dempster-Shafer belief and plausibility measures [44] on  $S\mathcal{L}$  which are not, in general, additive.

**Theorem 9.** *For all  $\theta, \varphi \in S\mathcal{L}$ , the following hold [33], [53]:*

- $\underline{\mu}(\theta) \leq \bar{\mu}(\theta)$

- $\underline{\mu}(\neg\theta) = 1 - \bar{\mu}(\theta)$  and  $\bar{\mu}(\neg\theta) = 1 - \underline{\mu}(\theta)$ .
- $\underline{\mu}(\theta \vee \varphi) = \underline{\mu}(\theta) + \underline{\mu}(\varphi) - \underline{\mu}(\theta \wedge \varphi)$  and  $\bar{\mu}(\theta \vee \varphi) = \bar{\mu}(\theta) + \bar{\mu}(\varphi) - \bar{\mu}(\theta \wedge \varphi)$

## 4 Truth Degrees as Mid-Points

We now introduce an additional type of uncertainty measure on  $S\mathcal{L}$  arising from the definition of a probability distribution on  $\mathbb{V}_k$ , which we will refer to as *truth degrees*. These are defined as the mid-point, or average, of the lower and upper Kleene belief measures given in definition 8.

**Definition 10.** *Truth Degrees*

Let  $(\underline{\mu}, \bar{\mu})$  be a Kleene belief pair on  $\mathcal{L}$  as given in definition 8, then the corresponding truth degree  $td : S\mathcal{L} \rightarrow [0, 1]$  is defined as the mid-point of  $\underline{\mu}$  and  $\bar{\mu}$  so that  $\forall \theta \in S\mathcal{L}$ :

$$td(\theta) = \frac{\underline{\mu}(\theta) + \bar{\mu}(\theta)}{2}$$

Notice that by definitions 8 and 10,  $td(\theta)$  corresponds to sum of the probability that  $\vec{v}(\theta) = \mathbf{t}$  and a half of the probability that  $\vec{v}(\theta) = \mathbf{b}$ ;

$$td(\theta) = w(\{\vec{v} : \vec{v}(\theta) = \mathbf{t}\}) + \frac{w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\})}{2} \text{ and}$$

$$td(\neg\theta) = w(\{\vec{v} : \vec{v}(\theta) = \mathbf{f}\}) + \frac{w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\})}{2}$$

In other words,  $td(\theta)$  is determined by reallocating the probability of  $\mathbf{b}$  evenly between  $\mathbf{t}$  and  $\mathbf{f}$ . Also, notice that by theorem 9 truth degrees satisfy the following two properties:  $\forall \theta, \varphi \in S\mathcal{L}$ ,

- $td(\neg\theta) = 1 - td(\theta)$
- $td(\theta \vee \varphi) = td(\theta) + td(\varphi) - td(\theta \wedge \varphi)$

These properties are also satisfied by probability measures on  $S\mathcal{L}$  (see Paris [40] for an exposition), but unlike probability, truth degrees do not always give values 0 and 1 to classical contradictions and tautologies<sup>6</sup>. In fact, it holds that:

$$td(\theta \wedge \neg\theta) = \frac{w(\{\vec{v} : \vec{v}(\theta \wedge \neg\theta) = \mathbf{b}\})}{2} = \frac{w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\})}{2} = \frac{\bar{\mu}(\theta) - \underline{\mu}(\theta)}{2}$$

Hence, if the agent has non-zero belief that  $\theta$  is a borderline sentence, i.e.  $\bar{\mu}(\theta) > \underline{\mu}(\theta)$ , then they will allocate non-zero truth degree to  $\theta \wedge \neg\theta$ , with an upper bound of 0.5 if they

---

<sup>6</sup>It is also the case that truth degree do not necessarily give the same value to classically equivalent sentences, but only to sentences which are equivalent in Kleene's logic

are certain of the borderline status of  $\theta$ . Similarly, if the agent is certain that  $\theta$  is **t**, i.e.  $\underline{\mu}(\theta) = \bar{\mu}(\theta) = 1$ , then  $td(\theta) = 1$ , whilst if they are certain that  $\theta$  is **f**, i.e.  $\underline{\mu}(\theta) = \bar{\mu}(\theta) = 0$ , then  $td(\theta) = 0$ . Hence, truth degree values of 0,  $\frac{1}{2}$  and 1 all correspond to states of total epistemic certainty.

It should be noted that we are somewhat reluctant to use the term *truth degree* for the mid-point measure in definition 10. In fact, our main motivation for doing so is to emphasise a link with fuzzy logic [55], the details of which we will describe below. However, to a certain extent the term is inappropriate since from definition 10 we can see that truth degrees are subjective measures of belief defined in the context of a three-valued truth model, and not truth values on a  $[0, 1]$  scale, as in the case of infinite valued logics [19]. This status as subjective belief will be made even more apparent when we consider a betting semantics for truth degrees in section 5.

Kleene belief pairs and truth degrees have been independently proposed by Williams [53] using a different but equivalent notation. Given a Kleene valuation, Williams proposes three distinct *loadings* in the form of binary and tertiary mappings defined on  $S\mathcal{L}$ ; these are Kleene loading corresponding to  $\underline{v}$ , LP loading corresponding to  $\bar{v}$  and symmetric loading corresponding to  $\frac{\underline{v} + \bar{v}}{2}$ <sup>7</sup>. Assuming a probability distribution on  $\mathbb{V}_k$ , [53] defines three measures on  $S\mathcal{L}$  by taking the expected value of the different loadings across  $\mathbb{V}_k$ , these then corresponding to  $\underline{\mu}$ ,  $\bar{\mu}$  and  $td$  respectively. In the light of this correspondence, [53] also states a version of theorem 9.

We now introduce *fuzzy truth degrees* as corresponding to a fully truth-functional (compositional) measure on  $S\mathcal{L}$ , with the same min-max combination rules as were originally proposed by Zadeh [55] for fuzzy membership functions. We will then show that fuzzy truth degrees are a special case of mid-point truth degrees as given in definition 10, but where the agent's only uncertainty concerns the correct level of vagueness (semantic precision) appropriate for the interpretation of  $\mathcal{L}$ .

**Definition 11.** *Fuzzy Truth Degree*

A fuzzy truth degree on  $\mathcal{L}$  is a function  $\zeta : S\mathcal{L} \rightarrow [0, 1]$  satisfying  $\forall \theta, \varphi \in S\mathcal{L}$ :

- $\zeta(\neg\theta) = 1 - \zeta(\theta)$
- $\zeta(\theta \wedge \varphi) = \min(\zeta(\theta), \zeta(\varphi))$
- $\zeta(\theta \vee \varphi) = \max(\zeta(\theta), \zeta(\varphi))$

**Theorem 12.** Let  $w$  be a probability distribution on  $\mathbb{V}_k$  such that  $\{\vec{v} : w(\vec{v}) > 0\} = \{\vec{v}_1, \dots, \vec{v}_m\}$  where  $\vec{v}_1 \preceq \dots \preceq \vec{v}_m$  and let  $(\underline{\mu}, \bar{\mu})$  be the associated Kleene belief pair on  $\mathcal{L}$ . Also, let  $td : S\mathcal{L} \rightarrow [0, 1]$  be the truth-degree generated from  $(\underline{\mu}, \bar{\mu})$  according to definition 10. Then in this case  $td$  is a fuzzy truth degree on  $\mathcal{L}$  as given in definition 11.

---

<sup>7</sup>LP stands for 'logic of paradox' and symmetric refers to symmetric logic (see Priest [42] for an overview).

*Proof.* We define binary functions  $b_i : S\mathcal{L} \rightarrow \{0, 1\}$  for  $i = 1, \dots, 2m$  such that:  $\forall \theta \in S\mathcal{L}$

$$b_i(\theta) = \begin{cases} \underline{v}_i(\theta) & : i \leq m \\ \bar{v}_{2m+1-i}(\theta) & : i > m \end{cases}$$

Now since  $\vec{v}_1 \preceq \dots \preceq \vec{v}_m$  it follows that  $b_1 \leq b_2 \leq \dots \leq b_{2m}$ . Also, if we define  $\forall \theta \in S\mathcal{L}$ ,  $i_\theta = \min\{i : b_i(\theta) = 1\}$  then  $\{b_i : b_i(\theta) = 1\} = \{b_i : i \geq i_\theta\}$ <sup>8</sup>. We now define a probability distribution  $w'$  on  $\{b_1, \dots, b_{2m}\}$  according to:

$$w'(b_i) = \begin{cases} \frac{w(\vec{v}_i)}{2} & : i \leq m \\ \frac{w(\vec{v}_{2m+1-i})}{2} & : i > m \end{cases}$$

From this we have that:

$$\begin{aligned} td(\theta) &= \frac{w(\{\vec{v} : \underline{v}(\theta) = 1\})}{2} + \frac{w(\{\vec{v} : \bar{v}(\theta) = 1\})}{2} \\ &= w'(b_i : i \leq m, b_i(\theta) = 1) + w'(b_i : i > m, b_i(\theta) = 1) \\ &= w'(b_i : b_i(\theta) = 1) = \sum_{i=i_\theta}^{2m} w'(b_i) \end{aligned}$$

Now for  $\theta, \varphi \in S\mathcal{L}$  we consider the following cases:

- $td(-\theta)$ : As already noted above, it follows trivially from definition 10 and theorem 9 that  $td(-\theta) = 1 - td(\theta)$ .
- $td(\theta \wedge \varphi)$ : By definition 2 we have that  $\forall i, b_i(\theta \wedge \varphi) = \min(b_i(\theta), b_i(\varphi))$ , hence  $i_{\theta \wedge \varphi} = \min\{i : b_i(\theta \wedge \varphi) = 1\} = \min\{i : \min(b_i(\theta), b_i(\varphi)) = 1\} = \min\{i : b_i(\theta) = 1 \text{ and } b_i(\varphi) = 1\} = \max(\min\{i : b_i(\theta) = 1\}, \min\{i : b_i(\varphi) = 1\}) = \max(i_\theta, i_\varphi)$ . Therefore,

$$\begin{aligned} td(\theta \wedge \varphi) &= \sum_{i=i_{\theta \wedge \varphi}}^{2m} w'(b_i) = \sum_{i=\max(i_\theta, i_\varphi)}^{2m} w'(b_i) \\ &= \min\left(\sum_{i=i_\theta}^{2m} w'(b_i), \sum_{i=i_\varphi}^{2m} w'(b_i)\right) = \min(td(\theta), td(\varphi)) \end{aligned}$$

- $td(\theta \vee \varphi)$ : By definition 2 we have that  $\forall i, b_i(\theta \vee \varphi) = \max(b_i(\theta), b_i(\varphi))$ , hence  $i_{\theta \vee \varphi} = \min(i_\theta, i_\varphi)$ . Therefore,

$$\begin{aligned} td(\theta \vee \varphi) &= \sum_{i=i_{\theta \vee \varphi}}^{2m} w'(b_i) = \sum_{i=\min(i_\theta, i_\varphi)}^{2m} w'(b_i) \\ &= \max\left(\sum_{i=i_\theta}^{2m} w'(b_i), \sum_{i=i_\varphi}^{2m} w'(b_i)\right) = \max(td(\theta), td(\varphi)) \end{aligned}$$

---

<sup>8</sup>In the case that  $\{b_i : b_i(\theta) = 1\} = \emptyset$  then for notational convenience we take,  $i_\theta = 2m+1$ . Subsequently, we then apply the convention that if  $j > 2m$  then  $\sum_{i=j}^{2m} w'(b_i) = 0$

□

**Theorem 13.** Let  $td : S\mathcal{L} \rightarrow [0, 1]$  be a fuzzy truth degree on  $\mathcal{L}$  generated as in theorem 12. Then  $\forall \theta \in S\mathcal{L}$ ;

$$\underline{\mu}(\theta) = \max(0, 2td(\theta) - 1) \text{ and } \bar{\mu}(\theta) = \min(1, 2td(\theta))$$

*Proof.* For  $\theta \in S\mathcal{L}$  we consider the following two cases:

- $td(\theta) \leq \frac{1}{2}$ : Now notice that by definition of  $w'$  it holds that  $w'(\{b_{m+1}, \dots, b_{2m}\}) = \frac{1}{2}$ . Hence, since  $td(\theta) \leq \frac{1}{2}$  then it follows that  $i_\theta \geq m + 1$ . Therefore;

$$\begin{aligned} \bar{\mu}(\theta) &= w(\{\vec{v} : \bar{v}(\theta) = 1\}) = 2w'(\{b_i : i > m, b_i(\theta) = 1\}) \\ &= 2w'(\{b_i : b_i(\theta) = 1\}) \text{ since } i_\theta \geq m + 1 = 2td(\theta) \end{aligned}$$

- $td(\theta) > \frac{1}{2}$ : In this case  $i_\theta \leq m$  by the above argument. From this it follows that:

$$b_i(\theta) = 1, \forall i > m \Rightarrow \bar{v}_i(\theta) = 1 \text{ for } i = 1, \dots, m \Rightarrow \bar{\mu}(\theta) = 1$$

Hence,  $\forall \theta \in S\mathcal{L}$ ,  $\bar{\mu}(\theta) = \min(1, 2td(\theta))$  as required. Then by duality:

$$\begin{aligned} \underline{\mu}(\theta) &= 1 - \bar{\mu}(\neg\theta) = 1 - \min(1, 2td(\neg\theta)) = 1 - \min(1, 2(1 - td(\theta))) \\ &= \max(0, 2td(\theta) - 1) \end{aligned}$$

□

The following corollary first appeared as a theorem in Lawry and Gonzalez [29], where it has a more direct proof. It is natural to include it here, however, as a consequence of theorems 12 and 13. As already noted elsewhere, e.g. in [29] and [31], it provides a clear bridge between Kleene belief pairs and interval-valued fuzzy logic [56].

**Corollary 14.** Let  $w$  be a probability distribution on  $\mathbb{V}_k$  such that  $\{\vec{v} : w(\vec{v}) > 0\} = \{\vec{v}_1, \dots, \vec{v}_m\}$  where  $\vec{v}_1 \preceq \dots \preceq \vec{v}_m$  and let  $(\underline{\mu}, \bar{\mu})$  be the associated Kleene belief pair on  $\mathcal{L}$ . Then  $\forall \theta, \varphi \in S\mathcal{L}$ ;

$$\begin{aligned} \underline{\mu}(\theta \wedge \varphi) &= \min(\underline{\mu}(\theta), \underline{\mu}(\varphi)), \quad \bar{\mu}(\theta \wedge \varphi) = \min(\bar{\mu}(\theta), \bar{\mu}(\varphi)) \text{ and} \\ \underline{\mu}(\theta \vee \varphi) &= \max(\underline{\mu}(\theta), \underline{\mu}(\varphi)), \quad \bar{\mu}(\theta \vee \varphi) = \max(\bar{\mu}(\theta), \bar{\mu}(\varphi)) \end{aligned}$$

*Proof.* By theorems 12 and 13 we have that  $\forall \theta, \varphi \in S\mathcal{L}$ ;

$$\begin{aligned} \underline{\mu}(\theta \wedge \varphi) &= \max(0, 2td(\theta \wedge \varphi) - 1) = \max(0, 2\min(td(\theta), td(\varphi)) - 1) \\ &= \min(\max(0, 2td(\theta) - 1), \max(0, 2td(\varphi) - 1)) = \min(\underline{\mu}(\theta), \underline{\mu}(\varphi)) \end{aligned}$$

The remaining cases then follow similarly. □

In the context of theorem 12 we can, in fact, now say something stronger about the relationship between Kleene belief measures and min-max fuzzy logic. Not only is it true that a special case of mid-point truth degrees (definition 10) are fuzzy truth degrees (definition 11), but also as we will show below, mid-point truth degrees which have  $w$  non-zero only on a sequence of increasingly sharp valuations, actually provide a complete characterisation of compositional fuzzy truth degrees.

**Lemma 15.** *Let  $\zeta : S\mathcal{L} \rightarrow [0, 1]$  be a fuzzy truth degree, as given in definition 11, and let  $w$  be a probability distribution on  $\mathbb{V}_k$  such that  $\{\vec{v} : w(\vec{v}) > 0\} = \{\vec{v}_1, \dots, \vec{v}_m\}$  where  $\vec{v}_1 \preceq \dots \preceq \vec{v}_m$ . Then  $\forall \theta \in S\mathcal{L}$ ,  $\zeta(\theta) = td(\theta)$  if and only if  $\forall l \in L\mathcal{L}$  (i.e. for all literals of  $\mathcal{L}$ ),  $\underline{\mu}(l) = \max(0, 2\zeta(l) - 1)$ .*

*Proof.*  $(\Rightarrow)$  if  $\forall \theta \in S\mathcal{L}$ ,  $td(\theta) = \zeta(\theta)$  then it follows trivially that  $\forall l \in L\mathcal{L}$ ,  $\underline{\mu}(l) = \max(0, 2\zeta(l) - 1)$  by theorem 13.

$(\Leftarrow)$   $\forall p \in \mathcal{P}$  we consider two cases:

- 1)  $\zeta(p) \neq \frac{1}{2}$ : In this case, then by definition 11 either  $\zeta(p) > \frac{1}{2}$  or  $\zeta(\neg p) > \frac{1}{2}$ . Let  $l \in \{p, \neg p\}$  be such that  $\zeta(l) > \frac{1}{2}$ . Now  $\underline{\mu}(l) = \max(0, 2\zeta(l) - 1) = 2\zeta(l) - 1 > 0$ . Also, by theorem 13  $\underline{\mu}(l) = \max(0, 2td(l) - 1)$ . Hence,  $\max(0, 2td(l) - 1) = 2\zeta(l) - 1 \Rightarrow 2td(l) - 1 = 2\zeta(l) - 1 \Rightarrow td(l) = \zeta(l) \Rightarrow td(\neg l) = \zeta(\neg l)$  by definition 11 and the general properties of mid-point truth degrees as described in the first part of this section. Hence,  $td(p) = \zeta(p)$  and  $td(\neg p) = \zeta(\neg p)$ .
- 2)  $\zeta(p) = \zeta(\neg p) = \frac{1}{2}$ : In this case  $\underline{\mu}(p) = \max(0, 2\zeta(p) - 1) = 0$ . Also, by theorem 13 we have that  $\underline{\mu}(p) = \max(0, 2td(p) - 1)$ . Therefore,  $\max(0, 2td(p) - 1) = 0 \Rightarrow td(p) \leq \frac{1}{2}$ . Similarly,  $\underline{\mu}(\neg p) = \max(0, 2\zeta(\neg p) - 1) = 0$  and by theorem 13 we have that  $\underline{\mu}(\neg p) = \max(0, 2td(\neg p) - 1)$ . Therefore,  $\max(0, 2td(\neg p) - 1) = 0 \Rightarrow td(\neg p) \leq \frac{1}{2}$ . Hence, since  $td(\neg p) = 1 - td(p)$  it follows immediately that  $td(p) = td(\neg p) = \frac{1}{2}$ .

From this argument we have that  $\forall l \in L\mathcal{L}$ ,  $td(l) = \zeta(l)$ . Hence, given the truth-functionality of both  $td$  and  $\zeta$  following from theorem 12 and definition 11 respectively, we have that  $\forall \theta \in S\mathcal{L}$ ,  $td(\theta) = \zeta(\theta)$ , as required.  $\square$

**Theorem 16.** *For any fuzzy truth degree  $\zeta : S\mathcal{L} \rightarrow [0, 1]$  on  $\mathcal{L}$ , as given in definition 11, there is a unique sequence  $\vec{v}_1 \preceq \dots \preceq \vec{v}_m$  of Kleene valuation pairs on  $\mathcal{L}$  and an associated probability distribution  $w$  on  $\mathbb{V}_k$  for which  $\{\vec{v} : w(\vec{v}) > 0\} = \{\vec{v}_1, \dots, \vec{v}_m\}$ , such that  $\forall \theta \in S\mathcal{L}$ ;*

$$\zeta(\theta) = td(\theta) = \frac{\underline{\mu}(\theta) + \overline{\mu}(\theta)}{2}$$

*Proof.* Using the orthopair notation for valuation pairs, then given any pair  $(P, N)$  we can naturally generate a set of literals  $F = P \cup \{\neg p : p \in N\}$ . Since  $(P, N)$  is an orthopair



so that  $P \cap N = \emptyset$ , it immediately follows that  $\forall p \in \mathcal{P}$ ,  $\{p, \neg p\} \not\subseteq F$ . Furthermore, given only  $F$  we can identify  $(P, N)$  by taking  $P = \{p : p \in F\}$  and  $N = \{p : \neg p \in F\}$ . Hence, the set  $\mathcal{F} = \{F \subseteq L\mathcal{L} : \forall p \in \mathcal{P}, \{p, \neg p\} \not\subseteq F\}$  of subsets of literals, provides an alternative characterisation of the set of orthopairs and consequently, by theorem 4, of  $\mathbb{V}_k$ . Also, note that by theorem 4, it follows that  $\forall l \in L\mathcal{L}$ ,  $\underline{v}(l) = 1$  if and only if  $l \in F$ .

By the above argument and theorem 7 it follows that any probability distribution  $w$  satisfying the required properties (i.e non-zero only on a sequence of valuations totally ordered by  $\preceq$ ) is characterised by a nested sequence of subsets of literals  $\{F_i : i\}$  such that  $F_i \in \mathcal{F}$  and  $F_i \subseteq F_{i+1}$ , together with weights  $w_i \in [0, 1]$  for which  $\sum_i w_i = 1$ . More specifically,  $F_i$  is a characterisation of valuation pair  $\vec{v}_i$  and  $w_i = w(\vec{v}_i)$ . Furthermore, given such a characterisation, we have that  $\forall l \in L\mathcal{L}$ ;

$$\underline{\mu}(l) = \sum_{F_i : l \in F_i} w_i$$

Notice that a sequence of this form together with an associated set of weights defines a consonant (or nested) random set on  $\mathcal{F}$  and, by the above equation,  $\underline{\mu} : L\mathcal{L} \rightarrow [0, 1]$  is then the corresponding single point coverage function.

Now let  $\{l : \zeta(l) > \frac{1}{2}\} = \{l_1, \dots, l_{m-1}\}$  ordered such that  $\zeta(l_{i-1}) \geq \zeta(l_i)$  for  $i = 2, \dots, m-1$ . Since  $\zeta(l) > \frac{1}{2} \Rightarrow \zeta(-l) < \frac{1}{2}$  it follows that the set  $\{l_1, \dots, l_{m-1}\} \in \mathcal{F}$  as are all its subsets. Given the above characterisation and lemma 15 it follows that the result holds if and only if there exists a nested sequence  $\{F_i : i\}$  where  $F_i \subseteq F_{i+1} \subseteq \{l_1, \dots, l_{m-1}\}$  with associated weights  $w_i$  where  $\sum_i w_i = 1$  and for which

$$\sum_{F_i : l_j \in F_i} w_i = 2\zeta(l_j) - 1 \text{ for } j = 1, \dots, m-1$$

Now from a well-known result according to which a consonant random set can be recovered from its single point coverage function (see theorem 2 in [18] or alternatively [24] and [8] for a more straightforward treatment of the finite case), it follows that the above set of equations has a solution given by:

$$F_m = \{l_1, \dots, l_{m-1}\}, \dots, F_i = \{l_1, \dots, l_{i-1}\}, \dots, F_2 = \{l_1\}, F_1 = \emptyset$$

and

$$w_m = 2\zeta(l_m) - 1, \dots, w_i = (2\zeta(l_{i-1}) - 1) - (2\zeta(l_i) - 1) = 2(\zeta(l_{i-1}) - \zeta(l_i)) \\ , \dots, w_1 = 1 - (2\zeta(l_1) - 1) = 2(1 - \zeta(l_1))$$

Furthermore, this solution is known to be unique once terms where  $w_i = 0$  are removed. The corresponding orthopairs representation can then be recovered as outlined above.<sup>9</sup>  $\square$

---

<sup>9</sup>It is possible that  $w_i = 0$  for some  $i \in \{1, \dots, m\}$ . For example, this could occur if  $\zeta(l_i) = \zeta(l_{i-1})$  or if  $\zeta(l_1) = 1$ . In this case, the resulting distribution  $w$  will be such that the length of the sequence of valuation pairs for which  $w(\vec{v}) > 0$  will be strictly less than  $m$  (rather than equal to  $m$ ). However, since in the statement of the theorem  $m$  is simply a variable taking integer values greater than or equal to 1, then this does not effect the result.

The proof of theorem 16 includes an algorithm for determining  $w$  and  $\vec{v}_1 \preceq \dots \preceq \vec{v}_m$  given fuzzy truth degree values on the propositional variables of  $\mathcal{L}$ . This algorithm is now illustrated in the following example.

**Example 17.** Let  $\mathcal{L}$  have propositional variables  $\mathcal{P} = \{p_1, p_2, p_3, p_4, p_5, p_6\}$ . Let  $\zeta$  be a fuzzy truth-degree on  $\mathcal{L}$  for which:

$$\zeta(p_1) = 0.6, \zeta(p_2) = 0.7, \zeta(p_3) = 0.85, \zeta(p_4) = 0.1, \zeta(p_5) = 0.2, \zeta(p_6) = 0.35$$

Now considering those literals  $l$  for which  $\zeta(l) > \frac{1}{2}$  we have:

$$\zeta(p_1) = 0.6, \zeta(p_2) = 0.7, \zeta(p_3) = 0.85, \zeta(\neg p_4) = 0.9, \zeta(\neg p_5) = 0.8, \zeta(\neg p_6) = 0.65$$

Resulting in the ordering:

$$\zeta(\neg p_4) > \zeta(p_3) > \zeta(\neg p_5) > \zeta(p_2) > \zeta(\neg p_6) > \zeta(p_1)$$

This gives us the following sequence of orthopairs with associated probabilities:

$$F_7 = \{\neg p_4, p_3, \neg p_5, p_2, \neg p_6, p_1\} \mapsto (P_7, N_7) = (\{p_3, p_2, p_1\}, \{p_4, p_5, p_6\}) : w_7 = 2(0.6) - 1 = 0.2$$

$$F_6 = \{\neg p_4, p_3, \neg p_5, p_2, \neg p_6\} \mapsto (P_6, N_6) = (\{p_3, p_2\}, \{p_4, p_5, p_6\}) : w_6 = 2(0.65 - 0.6) = 0.1$$

$$F_5 = \{\neg p_4, p_3, \neg p_5, p_2\} \mapsto (P_5, N_5) = (\{p_3, p_2\}, \{p_4, p_5\}) : w_5 = 2(0.7 - 0.65) = 0.1$$

$$F_4 = \{\neg p_4, p_3, \neg p_5\} \mapsto (P_4, N_4) = (\{p_3\}, \{p_4, p_5\}) : w_4 = 2(0.8 - 0.7) = 0.2$$

$$F_3 = \{\neg p_4, p_3\} \mapsto (P_3, N_3) = (\{p_3\}, \{p_4\}) : w_3 = 2(0.85 - 0.8) = 0.1$$

$$F_2 = \{\neg p_4\} \mapsto (P_2, N_2) = (\emptyset, \{p_4\}) : w_2 = 2(0.9 - 0.85) = 0.1$$

$$F_1 = \emptyset \mapsto (P_1, N_1) = (\emptyset, \emptyset) : w_1 = 2(1 - 0.9) = 0.2$$

Notice that  $(P_7, N_7)$  corresponds to a classical valuation since  $N_7 = P_7^c$ , whilst  $(P_1, N_1)$  corresponds to a fully vague model in which all sentences are borderline cases.

The general relationship between fuzzy truth degrees and the probabilities of  $\mathbf{t}$ ,  $\mathbf{b}$  and  $\mathbf{f}$  can be summarised as shown in figure 3. More specifically, the probability of each of the three truth values is a (piecewise) linear function of fuzzy truth degree as follows<sup>10</sup>:

- For  $td(\theta) \leq \frac{1}{2}$ : In this case  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{t}\}) = 0$ ,  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\}) = 2td(\theta)$ , and  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{f}\}) = 1 - 2td(\theta)$
- For  $td(\theta) > \frac{1}{2}$ : In this case  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{t}\}) = 2td(\theta) - 1$ ,  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\}) = 2(1 - td(\theta))$  and  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{f}\}) = 0$

<sup>10</sup>Given theorems 12 and 16 showing that fuzzy truth degrees are a special case of mid-point truth degrees, we will from now on also use the notation  $td$  when discussing the former.

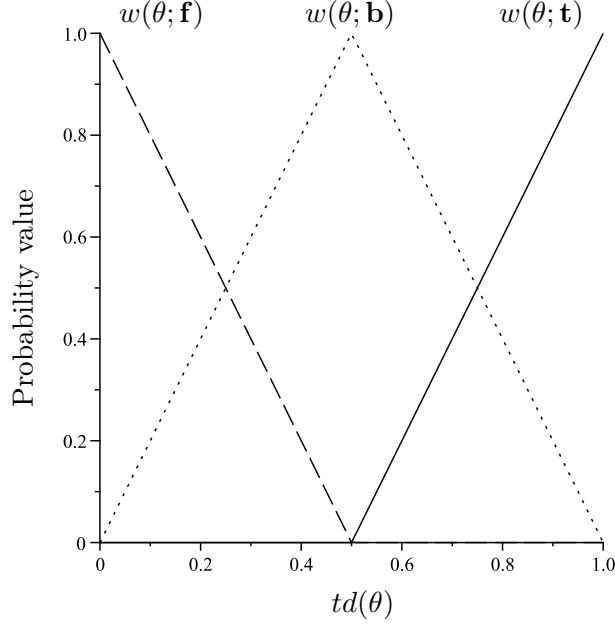


Figure 3: Plot showing the relationship between the fuzzy truth degree  $td(\theta)$  and the probabilities  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{f}\})$  (denoted  $w(\theta; \mathbf{f})$ ),  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\})$  (denoted  $w(\theta; \mathbf{b})$ ) and  $w(\{\vec{v} : \vec{v}(\theta) = \mathbf{t}\})$  (denoted  $w(\theta; \mathbf{t})$ )

In other words, if  $td(\theta) \in (0, 0.5)$  then it is certain that  $\theta$  is not absolutely true, but uncertain whether  $\theta$  is borderline or absolutely false. Similarly, if  $td(\theta) \in (0.5, 1)$  then it is certain that  $\theta$  is not absolutely false, but uncertain whether  $\theta$  is borderline or absolutely true. Given this we can begin to address issues like that of the police officer wondering what she can infer from a witness asserting that ‘the suspect is short’ with truth degree 0.6. We now see from figure 3 that such an assertion can be interpreted as the witness believing the suspect to be *absolutely short* with probability 0.2, and *borderline short* with probability 0.8.

## 5 A Betting Semantics for Truth Degrees

In this section we discuss a possible betting semantics for Kleene belief pairs and truth degrees taking as the point of departure de Finetti’s operational semantics for subjective probability [7]. By way of background we first give an overview of Paris’ generalization of de Finetti’s framework and state the main result from [41]. We will then exploit this result when considering bets under a three-valued truth model.

Let  $\mathbb{B}$  be a finite set of binary functions from  $S\mathcal{L}$  into  $\{0, 1\}$  representing the set of possible truth states. In de Finetti’s original formulation  $\mathbb{B}$  was the set of Tarski valuations on  $\mathcal{L}$ . However, in Paris’ generalization it is not restricted in this way. A bet on sentence  $\theta \in S\mathcal{L}$ , requiring stake  $s \in \mathbb{R}$ , and with odds  $\alpha \in [0, 1]$ , denoted by  $(s, \alpha, \theta)$ , is then

defined as follows:

- pay  $s \times \alpha$  pounds.
- if  $b(\theta) = 1$  then receive  $s$  pounds.
- if  $b(\theta) = 0$  then receive 0 pounds.

where  $b \in \mathbb{B}$  corresponds to the true state of the world. Notice that we permit  $s < 0$  which is interpreted as the agent selling rather than buying the bet. Now suppose an agent accepts a set of bets  $(s_i, \alpha_i, \theta_i)$  for  $i = 1, \dots, t$ , then her overall gain will be:

$$\sum_{i=1}^t s_i(b(\theta_i) - \alpha_i)$$

Furthermore, this set of bets is referred to as a *Dutch book* if the above expression is negative for every function  $b \in \mathbb{B}$ . In other words, a Dutch book is a set of bets which, if accepted, would result in a sure loss no matter what the true state of the world turns out to be.

We now define an agent's subjective belief in a sentence  $\theta$ , denoted  $\mathcal{B}(\theta)$ , to be the odds for which she will accept the bet  $(s, \mathcal{B}(\theta), \theta)$  for any stake  $s \in \mathbb{R}$ . In the light of this definition Paris [41] proves the following result:

**Theorem 18.** *There is no Dutch book acceptable to an agent adopting belief measure  $\mathcal{B}$ , if and only if there exists a probability distribution  $w$  on  $\mathbb{B}$  such that  $\forall \theta \in S\mathcal{L}$ ;*

$$\mathcal{B}(\theta) = w(\{b : b(\theta) = 1\})$$

Notice that if  $\mathbb{B}$  is the set of Tarski valuations on  $\mathcal{L}$  then theorem 18 means that in order to avoid Dutch bets,  $\mathcal{B}$  must be a probability measure on  $S\mathcal{L}$ .

Having set the scene we can now return to consider bets assuming an underlying three-valued truth model. In order to adapt the generalized de Finetti approach to this context we must decide how to deal with borderline outcomes. In other words, how is the bet  $(s, \alpha, \theta)$  decided if the truth value of  $\theta$  is **b**? One approach would be to agree that the agent only wins  $s$  if  $\theta$  is absolutely true and wins 0 otherwise. These type of bets, termed *lower bets*, are discussed in detail in Lawry and Tang [31]. Notice that by taking  $\mathbb{B} = \{\underline{v} : \vec{v} \in \mathbb{V}_k\}$  and applying theorem 18, it follows that in order for an agent to avoid Dutch books consisting of lower bets then she must adopt Kleene lower belief measures on  $S\mathcal{L}$  as her measures of subjective belief. There is then also a dual notion of *upper bets* in which a win occurs if and only if  $\theta$  is not absolutely false. This time taking  $\mathbb{B} = \{\bar{v} : \vec{v} \in \mathbb{V}_k\}$ , theorem 18 then forces a rational agent to adopt Kleene upper belief measures.

In order to provide a betting interpretation of truth degrees we now propose an alternative approach to dealing with borderline outcomes, which we will refer to as *decider bets*. The essential idea behind these bets is that winning or loosing in the case of a borderline outcome is determined on the basis of the truth-value of some fixed and inherently crisp proposition  $q \in \mathcal{P}$  i.e. this being the decider. In other words, a bet  $(s, \alpha, \theta)$  now takes the following form:

- Pay  $s \times \alpha$  pounds.
- If  $\vec{v}(\theta) = \mathbf{t}$  then receive  $s$  pounds.
- If  $\vec{v}(\theta) = \mathbf{f}$  then receive 0 pounds.
- If  $\vec{v}(\theta) = \mathbf{b}$  then,
  - if  $\vec{v}(q) = \mathbf{t}$  then receive  $s$  pounds, else
  - if  $\vec{v}(q) = \mathbf{f}$  then receive 0 pounds.

where  $\vec{v} \in \mathbb{V}_k(q)$  and  $\mathbb{V}_k(q) = \{\vec{v} \in \mathbb{V}_k : \vec{v}(q) \neq \mathbf{b}\}$  is the set of all possible truth states .

Decider bets can then be reformulated in terms of a certain class of binary functions  $\mathbb{B}$  defined as follows:  $\forall \vec{v} \in \mathbb{V}_k(q)$ , let  $b_{\vec{v}} : S\mathcal{L} \rightarrow \{0, 1\}$  be such that  $\forall \theta \in S\mathcal{L}$ ;

$$b_{\vec{v}}(\theta) = \begin{cases} 1 : \vec{v}(\theta) = \mathbf{t} \text{ or } (\vec{v}(\theta) = \mathbf{b} \text{ and } \vec{v}(q) = \mathbf{t}) \\ 0 : \text{otherwise} \end{cases}$$

Alternatively, taking  $v(q) = \underline{v}(q) = \bar{v}(q)$  then:

$$b_{\vec{v}}(\theta) = \max(\underline{v}(\theta), \min(\bar{v}(\theta), v(q)))$$

We then take:

$$\mathbb{B} = \{b_{\vec{v}} : \vec{v} \in \mathbb{V}_k(q)\}$$

Hence, the true state of the world generates a particular element  $b \in \mathbb{B}$  and we can then reformulate the decider bets so that they now fit within Paris's formulation. Consequently, a rational agent must define her belief measure  $\mathcal{B}$  in terms of a probability distribution  $w$  on this particular set of binary functions  $\mathbb{B}$  as given in theorem 18. The following result now shows that each valuation pair in  $\mathbb{V}_k(q)$  identifies a unique element of  $\mathbb{B}$ .

**Theorem 19.** *For  $\vec{v}_1, \vec{v}_2 \in \mathbb{V}_k(q)$  if  $\vec{v}_1 \neq \vec{v}_2$  then  $b_{\vec{v}_1} \neq b_{\vec{v}_2}$*

*Proof.* Notice that  $b_{\vec{v}}(q) = v(q)$  and hence we need only consider the case where  $\vec{v}_1(q) = \vec{v}_2(q)$  since otherwise the result follows trivially i.e. by taking  $\theta = q$ .

Suppose that  $\vec{v}_1(q) = \vec{v}_2(q) = \mathbf{t}$  and w.l.o.g consider the following cases:  $\exists \theta \in S\mathcal{L}$  such that

- $\vec{v}_1(\theta) = \mathbf{t}$  and  $\vec{v}_2(\theta) = \mathbf{f}$ : In this case  $b_{\vec{v}_1}(\theta) = 1$  and  $b_{\vec{v}_2}(\theta) = 0$ . Hence,  $b_{\vec{v}_1} \neq b_{\vec{v}_2}$ .
- $\vec{v}_1(\theta) = \mathbf{f}$  and  $\vec{v}_2(\theta) = \mathbf{b}$ : In this case  $b_{\vec{v}_1}(\theta) = 0$  and  $b_{\vec{v}_2}(\theta) = 1$ . Hence,  $b_{\vec{v}_1} \neq b_{\vec{v}_2}$ .
- $\vec{v}_1(\theta) = \mathbf{t}$  and  $\vec{v}_2(\theta) = \mathbf{b}$ : Notice in this case that  $b_{\vec{v}_1}(\theta) = b_{\vec{v}_2}(\theta) = 1$ . However, by duality  $\vec{v}_1(\neg\theta) = \mathbf{f}$  and  $\vec{v}_2(\neg\theta) = \mathbf{b}$  and hence from above  $b_{\vec{v}_1}(\neg\theta) = 0$  and  $b_{\vec{v}_2}(\neg\theta) = 1$ . Therefore,  $b_{\vec{v}_1} \neq b_{\vec{v}_2}$ .

Further suppose that  $\vec{v}_1(q) = \vec{v}_2(q) = \mathbf{f}$  and w.l.o.g consider the following cases:  $\exists \theta \in S\mathcal{L}$  such that

- $\vec{v}_1(\theta) = \mathbf{t}$  and  $\vec{v}_2(\theta) = \mathbf{f}$ : In this case  $b_{\vec{v}_1}(\theta) = 1$  and  $b_{\vec{v}_2}(\theta) = 0$ . Hence,  $b_{\vec{v}_1} \neq b_{\vec{v}_2}$ .
- $\vec{v}_1(\theta) = \mathbf{t}$  and  $\vec{v}_2(\theta) = \mathbf{b}$ : In this case  $b_{\vec{v}_1}(\theta) = 1$  and  $b_{\vec{v}_2}(\theta) = 0$ . Hence,  $b_{\vec{v}_1} \neq b_{\vec{v}_2}$ .
- $\vec{v}_1(\theta) = \mathbf{f}$  and  $\vec{v}_2(\theta) = \mathbf{b}$ : Notice in this case that  $b_{\vec{v}_1}(\theta) = b_{\vec{v}_2}(\theta) = 0$ . However, by duality  $\vec{v}_1(\neg\theta) = \mathbf{t}$  and  $\vec{v}_2(\neg\theta) = \mathbf{b}$  and hence from above  $b_{\vec{v}_1}(\neg\theta) = 1$  and  $b_{\vec{v}_2}(\neg\theta) = 0$ . Therefore,  $b_{\vec{v}_1} \neq b_{\vec{v}_2}$ .

□

From theorem 19 we see that  $w$  on  $\mathbb{B}$  naturally generates a probability distribution on  $\mathbb{V}_k(q)$ , also denoted  $w$ , such that  $w(\vec{v}) = w(b_{\vec{v}})$ . Hence, for the agent to avoid Dutch books consisting of decider bets,  $\mathcal{B}$  must satisfy that for all  $\theta \in S\mathcal{L}$ ;

$$\begin{aligned}\mathcal{B}(\theta) &= w(\{b_{\vec{v}} : b_{\vec{v}}(\theta) = 1\}) = w(\{\vec{v} : b_{\vec{v}}(\theta) = 1\}) \\ &= w(\{\vec{v} : \vec{v}(\theta) = \mathbf{t}\}) + w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}, \vec{v}(q) = \mathbf{t}\})\end{aligned}$$

Now for  $\mathcal{B}(\theta)$  to correspond to a truth degree value of  $\theta$  as given by definition 10 requires that;

$$w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}, \vec{v}(q) = \mathbf{t}\}) = w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}, \vec{v}(q) = \mathbf{f}\}) = \frac{w(\{\vec{v} : \vec{v}(\theta) = \mathbf{b}\})}{2}$$

One scenario in which this assumption could be valid is as follows: Let  $\mathcal{L}'$  denote  $\mathcal{L}$  restricted to the propositional variables  $\mathcal{P} - \{q\}$ , and  $S\mathcal{L}'$  be the sentences of  $\mathcal{L}'$ . Then we assume that  $q$  is chosen so as to be independent of the sentences in  $S\mathcal{L}'$  and furthermore, such that the agent's belief in  $q$  being true is  $\frac{1}{2}$ . By independent, we mean that for any sentence  $\theta \in S\mathcal{L}'$  the truth value of  $q$ , i.e.  $\vec{v}(q)$ , is independent of the truth value of  $\theta$ , i.e.  $\vec{v}(\theta)$ . Hence, we are assuming that  $w$  on  $\mathbb{V}_k(q)$  satisfies the following:  $\forall \theta \in S\mathcal{L}'$

$$w(\vec{v}(q)|\vec{v}(\theta)) = w(\vec{v}(q)) \text{ and } w(\vec{v}(q) = \mathbf{t}) = w(\vec{v}(q) = \mathbf{f}) = \frac{1}{2}$$

For example,  $q$  might refer to the outcome from tossing a certain fair coin as being a head, whilst  $\mathcal{P} - \{q\}$  could be propositions referring to characteristics of the next person, call them  $x$ , to walk through a door e.g.  $x$  is tall,  $x$  is handsome,  $x$  is blonde etc.

**Example 20.** Let  $\mathcal{L}$  be such that  $\mathcal{P} = \{p_1, p_2, p_3, p_4, q\}$  and let  $\mathbb{V}_k(q)$  be the set of all Kleene valuation pairs on  $\mathcal{L}$  for which  $\vec{v}(q) \neq \mathbf{b}$ . Let  $w$  be the probability distribution on  $\mathbb{V}_k(q)$  defined in orthopairs notation as follows:

$$\begin{aligned} &(\{p_1, q\}, \emptyset) : 0.1, (\{p_1\}, \{q\}) : 0.1, (\{p_1, p_2, q\}, \emptyset) : 0.15, (\{p_1, p_2\}, \{q\}) : 0.15, \\ &(\{p_1, p_2, q\}, \{p_3\}) : 0.2, (\{p_1, p_2\}, \{p_3, q\}) : 0.2, \\ &(\{p_1, p_2, q\}, \{p_3, p_4\}) : 0.05, (\{p_1, p_2\}, \{p_3, p_4, q\}) : 0.05 \end{aligned}$$

In this case all of the above conditions are satisfied. For example, consider  $\theta = p_2 \wedge \neg p_3 \in S\mathcal{L}'$ , then the following holds:

$$w(\vec{v}(q) = \mathbf{t} | \vec{v}(p_2 \wedge \neg p_3) = \mathbf{b}) = w(\vec{v}(q) = \mathbf{f} | \vec{v}(p_2 \wedge \neg p_3) = \mathbf{b}) = \frac{1}{2} = w(\vec{v}(q) = \mathbf{t})$$

To see this notice that the only valuation pairs with non-zero probability for which  $\vec{v}(p_2 \wedge \neg p_3) = \mathbf{b}$  are the following;

$$(\{p_1, q\}, \emptyset), (\{p_1\}, \{q\}), (\{p_1, p_2, q\}, \emptyset), (\{p_1, p_2\}, \{q\})$$

Amongst this subset of valuations the probability is then split evenly between those for which  $\vec{v}(q) = \mathbf{t}$  and those for which  $\vec{v}(q) = \mathbf{f}$ . Furthermore, notice that if we restrict ourselves to sentences in  $S\mathcal{L}'$  then the corresponding marginal distribution for valuation pairs on  $\mathcal{L}'$  is as follows:

$$(\{p_1\}, \emptyset) : 0.2, (\{p_1, p_2\}, \emptyset) : 0.3, (\{p_1, p_2\}, \{p_3\}) : 0.4, (\{p_1, p_2\}, \{p_3, p_4\}) : 0.1$$

Furthermore, note that the above valuation pairs with non-zero probability form a nested sequence as required in theorem 7 and hence the truth degree generated by  $w$  according to definition 10, corresponds to a fuzzy truth degree (definition 11) when restricted to  $S\mathcal{L}'$ .

## 6 Truth Degrees in a Classical or Supervaluation Framework

As discussed in section 2 Kleene valuations pairs are only one of several possible models to account of truth-gaps in a propositional language. Furthermore, the fact that, amongst other features, they permit borderline contradictions, makes them controversial. Indeed, we might also view such a liberal allocation of borderline status in the light of a more general failure to represent *penumbral connections*. Fine [16] introduced the notion of penumbral connections as corresponding to ‘logical relations [that] holds between indefinite sentences’. As can be seen from table 1 the Kleene three-valued truth tables are conservative, in that both the conjunction and the disjunction of two borderline sentences are *always* also borderline cases. This property ultimately means that many penumbral connections simply cannot be captured within the Kleene framework. On the other hand,

as discussed in [31], supervaluations are better able to represent absolute relationships between borderline sentences. In particular,  $\theta \wedge \neg\theta$  is always absolutely false in the supervaluationist model irrespective of the truth value of  $\theta$ . As noted in section 2, supervaluations on a propositional language can also be expressed within the valuation pair notation. In subsection 6.2 we will investigate supervaluation pairs together with their associated belief pairs and truth degrees, and describe their relationships to min-max fuzzy logic.

Before discussing supervaluations we will initially, in subsection 6.1, investigate fuzziness when assuming a purely classical (Tarskian) truth model. This is consistent with Lindley's [35] and Cheeseman's [3] claims that vagueness can be captured entirely within classical probability theory. From the perspective of the earlier discussion in section 1, in this case we will be equating vagueness with semantic uncertainty and denying the existence of explicit borderline cases. Instead, a borderline case of a predicate will simply be interpreted as one in which the probability of the predicate and its negation holding are both close to 0.5. In fact, the assumption of an underlying classical truth model is common to many of the proposed probabilistic semantics for fuzziness including [20], [27], [37], [46] and [48]. Here we will show that under certain circumstances the classical probabilistic version of truth degrees, i.e. probability measures on  $S\mathcal{L}$ , can be consistent with fuzzy truth degrees but only on a restricted subset of the sentences of  $\mathcal{L}$ .

## 6.1 Truth Degrees and Sequences of Tarski Valuations

Let  $\mathbb{V}_c$  denote the set of classical or Tarski valuations on  $\mathcal{L}$ . We also define  $S\mathcal{L}^+$  as the sentences of  $\mathcal{L}$  generated recursively from the propositional variables by application of the connectives  $\wedge$  and  $\vee$  only, and similarly we let  $S\mathcal{L}^-$  denote the sentences of  $\mathcal{L}$  generated recursively from the negated propositional variables by application of  $\wedge$  and  $\vee$  only. Here we can think of  $S\mathcal{L}^+$  as the set of *entirely positive* sentences and  $S\mathcal{L}^-$  as the set of *entirely negative* sentences of  $\mathcal{L}$  respectively. Note that the class of sentences  $S\mathcal{L}^+ \cup S\mathcal{L}^-$ , whilst restricted, is nonetheless important for applications e.g. data base querying and rule-based systems. We now introduce a partial ordering on  $\mathbb{V}_c$  which has some similarity to the semantic precision ordering on valuation pairs.

**Definition 21.** *Ordering on  $\mathbb{V}_c$*

For  $v_1, v_2 \in \mathbb{V}_c$ ,  $v_1 \trianglelefteq v_2$  iff  $\forall p \in \mathcal{P}$ ,  $v_1(p) = 1 \Rightarrow v_2(p) = 1$ .

**Theorem 22.** *For  $v_1, v_2 \in \mathbb{V}_c$ , if  $v_1 \trianglelefteq v_2$  then;*

$$(i) \quad \forall \psi \in S\mathcal{L}^+, v_1(\psi) \leq v_2(\psi).$$

$$(ii) \quad \forall \psi \in S\mathcal{L}^-, v_2(\psi) \leq v_1(\psi)$$

*Proof.* Part (i): We proceed by induction on the complexity of sentences in  $S\mathcal{L}^+$ . Let  $S\mathcal{L}^{+,0} = \mathcal{P}$  and  $S\mathcal{L}^{+,k} = S\mathcal{L}^{+,k-1} \cup \{\theta \wedge \varphi, \theta \vee \varphi : \theta, \varphi \in S\mathcal{L}^{+,k-1}\}$  for  $k \geq 1$ . Now if



$\psi = p_i \in S\mathcal{L}^{+,0} = \mathcal{P}$  then  $v_1(p_i) \leq v_2(p_i)$  by definition 21. If  $\psi \in S\mathcal{L}^{+,k}$  then either  $\psi \in S\mathcal{L}^{+,k-1}$  in which case the result follows trivially or there exist  $\theta, \varphi \in S\mathcal{L}^{+,k-1}$  and one of the following holds:

- $\psi = \theta \wedge \varphi$ : In this case  $v_1(\psi) = \min(v_1(\theta), v_1(\varphi)) \leq \min(v_2(\theta), v_2(\varphi))$  (by induction)  $= v_2(\psi)$ .
- $\psi = \theta \vee \varphi$ : In this case  $v_1(\psi) = \max(v_1(\theta), v_1(\varphi)) \leq \max(v_2(\theta), v_2(\varphi))$  (by induction)  $= v_2(\psi)$ .

Part (ii): Let  $S\mathcal{L}^{-,0} = \{\neg p_i : p_i \in \mathcal{P}\}$  and  $S\mathcal{L}^{-,k} = S\mathcal{L}^{-,k-1} \cup \{\theta \wedge \varphi, \theta \vee \varphi : \theta, \varphi \in S\mathcal{L}^{-,k-1}\}$  for  $k \geq 1$ . Now if  $\psi = \neg p_i \in S\mathcal{L}^{-,0}$  then  $v_2(\neg p_i) = 1 \Rightarrow v_2(p_i) = 0 \Rightarrow v_1(p_i) = 0$ , since  $v_1 \leq v_2, \Rightarrow v_1(\neg p_i) = 1$  as required. The inductive steps then mirror those of part (i).  $\square$

Hence, from definition 21 and theorem 22 we see if  $v_1 \leq v_2$  then  $v_2$  has a greater tendency than  $v_1$  to classify positive sentences as being true, and a lesser tendency to classify negative sentences as being true. This is in contrast to the semantic precision ordering on Kleene valuations, where  $\vec{v}_1 \preceq \vec{v}_2$  means that  $\vec{v}_2$  has a greater tendency that  $\vec{v}_1$  to classify *any* of the sentences of  $\mathcal{L}$  as being true.

**Definition 23.** Given  $v_1, \dots, v_m \in \mathbb{V}_c$  forming a sequence  $v_1 \leq v_2 \leq \dots \leq v_m$  then  $\forall \theta \in S\mathcal{L}$ , let  $l_\theta = \min\{i : v_i(\theta) = 1\}$  and let  $u_\theta = \max\{i : v_i(\theta) = 1\}$

**Theorem 24.** Given  $v_1, \dots, v_m \in \mathbb{V}_c$  forming a sequence  $v_1 \leq v_2 \leq \dots \leq v_m$  then:

- (i)  $\forall \theta, \varphi \in S\mathcal{L}^+, l_{\theta \wedge \varphi} = \max(l_\theta, l_\varphi)$  and  $l_{\theta \vee \varphi} = \min(l_\theta, l_\varphi)$
- (ii)  $\forall \theta, \varphi \in S\mathcal{L}^-, u_{\theta \wedge \varphi} = \min(u_\theta, u_\varphi)$  and  $u_{\theta \vee \varphi} = \max(u_\theta, u_\varphi)$

*Proof.* Part (i): By theorem 22 we have that  $\forall \theta \in S\mathcal{L}^+, \{i : v_i(\theta) = 1\} = \{l_\theta, \dots, m\}$  and hence,

$$\begin{aligned} \{i : v_i(\theta \wedge \varphi) = 1\} &= \{i : v_i(\theta) = 1\} \cap \{i : v_i(\varphi) = 1\} = \{l_\theta, \dots, m\} \cap \{l_\varphi, \dots, m\} \\ &= \{\max(l_\theta, l_\varphi), \dots, m\} \Rightarrow l_{\theta \wedge \varphi} = \max(l_\theta, l_\varphi) \end{aligned}$$

as required.

Similarly,

$$\begin{aligned} \{i : v_i(\theta \vee \varphi) = 1\} &= \{i : v_i(\theta) = 1\} \cup \{i : v_i(\varphi) = 1\} = \{l_\theta, \dots, m\} \cup \{l_\varphi, \dots, m\} \\ &= \{\min(l_\theta, l_\varphi), \dots, m\} \Rightarrow l_{\theta \vee \varphi} = \min(l_\theta, l_\varphi) \end{aligned}$$

as required.

Part (ii): By theorem 22 we have that  $\forall \theta \in S\mathcal{L}^-, \{i : v_i(\theta) = 1\} = \{1, \dots, u_\theta\}$ . Hence,

$$\begin{aligned} \{i : v_i(\theta \wedge \varphi) = 1\} &= \{i : v_i(\theta) = 1\} \cap \{i : v_i(\varphi) = 1\} = \{1, \dots, u_\theta\} \cap \{1, \dots, u_\varphi\} \\ &= \{1, \dots, \min(u_\theta, u_\varphi)\} \Rightarrow u_{\theta \wedge \varphi} = \min(u_\theta, u_\varphi) \end{aligned}$$

as required. Also,

$$\begin{aligned}\{i : v_i(\theta \vee \varphi) = 1\} &= \{i : v_i(\theta) = 1\} \cup \{i : v_i(\varphi) = 1\} = \{1, \dots, u_\theta\} \cup \{1, \dots, u_\varphi\} \\ &= \{1, \dots, \max(u_\theta, u_\varphi)\} \Rightarrow u_{\theta \vee \varphi} = \max(u_\theta, u_\varphi)\end{aligned}$$

as required.  $\square$

Since for classical valuations there are no borderline cases then given a probability distribution  $w$  on  $\mathbb{V}_c$ , truth degrees are simply defined such that  $\forall \theta \in S\mathcal{L}$ ;

$$td(\theta) = w(\{v : v(\theta) = 1\})$$

In this case  $td$  is just a probability measure on  $S\mathcal{L}$ .

**Theorem 25.** *Let  $w$  be a probability distribution on  $\mathbb{V}_c$  such that  $\{v : w(v) > 0\} = \{v_1, \dots, v_m\}$  where  $v_1 \trianglelefteq v_2 \trianglelefteq \dots \trianglelefteq v_m$ . Then  $\forall \theta, \varphi \in S\mathcal{L}^+(S\mathcal{L}^-)$  the following hold:*

$$\begin{aligned}td(\theta \wedge \varphi) &= \min(td(\theta), td(\varphi)) \text{ and } td(\theta \vee \varphi) = \max(td(\theta), td(\varphi)) \\ \text{where } td(\theta) &= w(\{v : v(\theta) = 1\})\end{aligned}$$

*Proof.* Consider  $\theta, \varphi \in S\mathcal{L}^+$  then by theorem 24 it follows that:

$$\begin{aligned}td(\theta \wedge \varphi) &= w(\{v_i : v_i(\theta \wedge \varphi) = 1\}) = \sum_{r=l_{\theta \wedge \varphi}}^m w(v_r) = \sum_{r=\max(l_\theta, l_\varphi)}^m w(v_r) \\ &= \min\left(\sum_{r=l_\theta}^m w(v_r), \sum_{r=l_\varphi}^m w(v_r)\right) = \min(td(\theta), td(\varphi))\end{aligned}$$

as required. Also,

$$\begin{aligned}td(\theta \vee \varphi) &= w(\{v_i : v_i(\theta \vee \varphi) = 1\}) = \sum_{r=l_{\theta \vee \varphi}}^m w(v_r) = \sum_{r=\min(l_\theta, l_\varphi)}^m w(v_r) \\ &= \max\left(\sum_{r=l_\theta}^m w(v_r), \sum_{r=l_\varphi}^m w(v_r)\right) = \max(td(\theta), td(\varphi))\end{aligned}$$

as required.

Now consider  $\theta, \varphi \in S\mathcal{L}^-$  the by theorem 24 we have that:

$$\begin{aligned}td(\theta \wedge \varphi) &= w(\{v_i : v_i(\theta \wedge \varphi) = 1\}) = \sum_{i=1}^{u_{\theta \wedge \varphi}} w(v_i) = \sum_{i=1}^{\min(u_\theta, u_\varphi)} w(v_i) \\ &= \min\left(\sum_{i=1}^{u_\theta} w(v_i), \sum_{i=1}^{u_\varphi} w(v_i)\right) = \min(td(\theta), td(\varphi))\end{aligned}$$

as required. Also,

$$\begin{aligned} td(\theta \vee \varphi) &= w(\{v_i : v_i(\theta \vee \varphi) = 1\}) = \sum_{i=1}^{u_{\theta \vee \varphi}} w(v_r) = \sum_{i=1}^{\max(u_\theta, u_\varphi)} w(v_r) \\ &= \max(\sum_{i=1}^{u_\theta} w(v_r), \sum_{i=1}^{u_\varphi} w(v_r)) = \max(td(\theta), td(\varphi)) \end{aligned}$$

as required.  $\square$

One example where we can find sequences of classical valuations as in theorem 25, relates to a prototype theory interpretation of categories as introduced in [28]. Suppose that the propositions  $\mathcal{P}$  correspond to the formula  $\{Q_1(x), \dots, Q_n(x)\}$  where  $\{Q_i : i = 1, \dots, n\}$  are unary predicates and  $x$  is particular example. Furthermore, suppose that the interpretation of these predicates is as follows: Given an underlying universe  $\Omega$  on which is defined a pseudo-distance metric  $d$ , then for any element  $x \in \Omega$ ,  $Q_i(x)$  holds if and only if  $d(x, a_i) \leq \epsilon$  where  $a_i \in \Omega$  is the prototype for  $Q_i$  and  $\epsilon \in \mathbb{R}^+$  is a distance threshold. Hence, for a fixed threshold value  $\epsilon$  this model naturally generates a classical valuation  $v_\epsilon$  on  $\mathcal{L}$  such that;  $\forall p_i \in \mathcal{P}$

$$v_\epsilon(p_i) = 1 \text{ if and only if } d(x, a_i) \leq \epsilon$$

In other words,  $p_i$  is true if and only if  $x$  is *sufficiently similar* to the prototype  $a_i$ . Hence, a sequence of increasing thresholds  $\epsilon_1 < \epsilon_2 < \dots < \epsilon_m$  naturally generates a sequence of classical valuations on  $\mathcal{L}$  such that  $v_{\epsilon_1} \leq v_{\epsilon_2} \leq \dots \leq v_{\epsilon_m}$ . One can then envisage a scenario in which an agent's uncertainty is only regarding the value of the threshold  $\epsilon$ . If they were then to define a probability distribution on  $\epsilon$  this would result in a corresponding distribution  $w$  on  $\mathbb{V}_c$ , non-zero only on a sequence of valuations as in theorem 25.

## 6.2 Truth Degrees and Supervaluation Pairs

Supervaluationism was proposed as a theory of vagueness by Fine [16] (see also Williamson [54] for an exposition). In this approach it is assumed that vague predicates have different admissible crisp interpretations, referred to as *precisifications*. For example, the predicate *short* may admit a range of admissible threshold values on height, each defining a different precisification. The simplest formulation of supervaluationism in a propositional logic language is in terms of a set of admissible classical valuations. This approach naturally leads to a valuation pair representation as follows:

**Definition 26.** *Supervaluation Pairs [31]*

A *supervaluation pair* is defined as follows: For  $\Pi \subseteq \mathbb{V}_c$  let  $\forall \theta \in S\mathcal{L}$ ;

$$\underline{v}(\theta) = \min\{v(\theta) : v \in \Pi\} \text{ and } \overline{v}(\theta) = \max\{v(\theta) : v \in \Pi\}$$

Let  $\mathbb{V}_s$  denote the set of all supervaluation pairs on  $\mathcal{L}$ .<sup>11</sup>

Lawry and Tang [31] show that supervaluation pairs are characterised by the following four properties:  $\forall \theta, \varphi \in S\mathcal{L}$ ;

- *Duality*:  $\underline{v}(\neg\theta) = 1 - \bar{v}(\theta)$  and  $\bar{v}(\neg\theta) = 1 - \underline{v}(\theta)$
- *Tautology preservation*: If  $\theta$  is a classical tautology then  $\vec{v}(\theta) = \mathbf{t}$ .
- *Equivalence*: If  $\theta$  and  $\varphi$  are classically equivalent then  $\vec{v}(\theta) = \vec{v}(\varphi)$ .
- *Maximum upper*:  $\bar{v}(\theta \vee \varphi) = \max(\bar{v}(\theta), \bar{v}(\varphi))$ .

Notice that by duality and tautology preservation any classical contradiction  $\theta \wedge \neg\theta$  is absolutely false no matter what the truth value of  $\theta$ . Also, by duality and maximal upper we have that:

$$\underline{v}(\theta \wedge \varphi) = \min(\underline{v}(\theta), \underline{v}(\varphi))$$

In general, however, for supervaluation pairs it only holds that:

$$\underline{v}(\theta \vee \varphi) \geq \max(\underline{v}(\theta), \underline{v}(\varphi)) \text{ and } \bar{v}(\theta \wedge \varphi) \leq \min(\bar{v}(\theta), \bar{v}(\varphi))$$

In comparison to the Kleene valuation pair operators (definition 2) this means that supervaluation pairs have a lesser tendency to propagate borderline truth values.

We now introduce a particular sub-class of supervaluation pairs on  $\mathcal{L}$  referred to as *bounded supervaluation pairs*.

**Definition 27.** *Bounding Valuations*

Let  $\vec{v} \in \mathbb{V}_s$  be a supervaluation pair characterised by the set of admissible classical valuations  $\Pi \subseteq \mathbb{V}_c$ . Then let  $v_*, v^* \in \mathbb{V}_c$  be classical valuations such that:

$$\forall p_i \in \mathcal{P}, v_*(p_i) = \min\{v(p_i) : v \in \Pi\} \text{ and } v^*(p_i) = \max\{v(p_i) : v \in \Pi\}$$

**Definition 28.** *Bounded Supervaluation pairs [31]*

$\vec{v} \in \mathbb{V}_s$  is a bounded supervaluation pair if and only if  $\{v_*, v^*\} \subseteq \Pi$ . Let  $\mathbb{V}_{bs}$  denote the set of all bounded supervaluation pairs on  $\mathcal{L}$ .

One way to motivate bounded supervaluation pairs is to consider disjunctions and conjunctions of the borderline propositions. For  $\vec{v} \in \mathbb{V}_s$  let  $B$  denote the set of borderline propositional variables i.e.  $B = \{p_i : \vec{v}(p_i) = \mathbf{b}\}$ . Now if  $v_* \notin \Pi$  then  $\vec{v}(\bigvee_{p_i \in B} p_i) = \mathbf{t}$  and if  $v^* \notin \Pi$  then  $\vec{v}(\bigwedge_{p_i \in B} p_i) = \mathbf{f}$ . These would seem to be rather strong properties

---

<sup>11</sup>For any supervaluation pair  $\vec{v} \in \mathbb{V}_s$  there is a *unique* defining set of admissible classical valuations given by  $\Pi = \{v \in \mathbb{V}_c : \bar{v}(\alpha_v) = 1\}$  where  $\alpha_v = \bigwedge_{p_i : v(p_i)=1} p_i \wedge \bigwedge_{p_i : v(p_i)=0} \neg p_i$ .

for borderline cases<sup>12</sup>. Hence, the restriction to bounded supervaluation pairs means that for the borderline propositions there will be at least one admissible valuation in which they are all true, and at least one in which they are all false. This ensures that  $\vec{v}(\bigvee_{p_i \in B} p_i) = \vec{v}(\bigwedge_{p_i \in B} p_i) = \mathbf{b}$ .

The following theorem from [31] shows that bounded supervaluation pairs obey all the Kleene valuation pairs conjunction and disjunction operators when restricted to either the set of entirely positive sentences  $S\mathcal{L}^+$  or the set of entirely negative sentences  $S\mathcal{L}^-$ . This result will be a foundation of the bridge we will build between the supervaluationist framework and fuzzy logic.

**Theorem 29.** *For  $\vec{v} \in \mathbb{V}_{bs}$  then  $\forall \theta, \varphi \in S\mathcal{L}^+(S\mathcal{L}^-)$  it holds that:*

$$\bar{v}(\theta \wedge \varphi) = \min(\bar{v}(\theta), \bar{v}(\varphi)) \text{ and } \underline{v}(\theta \vee \varphi) = \max(\underline{v}(\theta), \underline{v}(\varphi))$$

The notion of semantic precision is clearly also relevant in a supervaluationist framework and definition 6 can be naturally extended so as to generate a vagueness ordering, also denoted  $\preceq$ , on  $\mathbb{V}_s$ . In this case it immediately follows that for  $\vec{v}_1, \vec{v}_2 \in \mathbb{V}_s$ , if  $\vec{v}_1 \preceq \vec{v}_2$  then  $v_{1*} \leq v_{2*}$  and  $v_2^* \leq v_1^*$ , since by definition 27 it holds that  $\forall \vec{v} \in \mathbb{V}_s$  and  $\forall p_i \in \mathcal{P}$ ,  $v_*(p_i) = \underline{v}(p_i)$  and  $v^*(p_i) = \bar{v}(p_i)$ . Furthermore, we have the following characterisation of the  $\preceq$  ordering on  $\mathbb{V}_s$  taken from [31].

**Theorem 30.** [31] *For  $\vec{v}_1, \vec{v}_2 \in \mathbb{V}_s$ ,  $\vec{v}_1 \preceq \vec{v}_2$  if and only if  $\Pi_1 \subseteq \Pi_2$ .*

In the following theorem we now show that, for bounded supervaluation pairs, there is a close relationship between the lower and upper valuations and the bounding classical valuations (definition 27) on  $S\mathcal{L}^+$  and  $S\mathcal{L}^-$ .

**Theorem 31.** *If  $\vec{v} \in \mathbb{V}_{bs}$  then:*

- (i)  $\forall \psi \in S\mathcal{L}^+$ ,  $\underline{v}(\psi) = v_*(\psi)$  and  $\bar{v}(\psi) = v^*(\psi)$
- (ii)  $\forall \psi \in S\mathcal{L}^-$ ,  $\underline{v}(\psi) = v^*(\psi)$  and  $\bar{v}(\psi) = v_*(\psi)$

*Proof.* Part (i): Proceed by induction on the complexity of sentences. For  $\psi = p_i \in S\mathcal{L}^{+,0} = \mathcal{P}$  then by definition 27  $\underline{v}(p_i) = \min\{v(p_i) : v \in \Pi\} = v_*(p_i)$  and  $\bar{v}(p_i) = \max\{v(p_i) : v \in \Pi\} = v^*(p_i)$ . If  $\psi \in S\mathcal{L}^{+,k}$  then either  $\psi \in S\mathcal{L}^{+,k-1}$  in which case the result follows trivially or there exists  $\theta, \varphi \in S\mathcal{L}^{+,k-1}$  for which one of the following holds:

- $\psi = \theta \wedge \varphi$ : In this case  $\underline{v}(\psi) = \underline{v}(\theta \wedge \varphi) = \min(\underline{v}(\theta), \underline{v}(\varphi))$  by the properties of supervaluation pairs as outlined above [31]  $= \min(v_*(\theta), v_*(\varphi))$  by induction  $= v_*(\theta \wedge \varphi) =$

---

<sup>12</sup>As an example of an unbounded supervaluation pair assume  $\mathcal{P} = \{p_1, p_2\}$  and let  $\Pi = \{v_1, v_2\}$  where  $v_1(p_1 \wedge \neg p_2) = 1$  and  $v_2(\neg p_1 \wedge p_2) = 1$ . Then  $v_*$  is such that  $v_*(\neg p_1 \wedge \neg p_2) = 1$  and  $v^*$  is such that  $v^*(p_1 \wedge p_2) = 1$ . Clearly in this case  $\{v_*, v^*\} \cap \Pi = \emptyset$ .

$v_*(\psi)$  by the properties of Tarski valuations. Also  $\bar{v}(\psi) = \bar{v}(\theta \wedge \varphi) = \min(\bar{v}(\theta), \bar{v}(\varphi))$  by theorem 29  $= \min(v^*(\theta), v^*(\varphi))$  by induction  $= v^*(\theta \wedge \varphi) = v^*(\psi)$  by the properties of Tarski valuations.

- $\psi = \theta \vee \varphi$ : In this case  $\underline{v}(\psi) = \underline{v}(\theta \vee \varphi) = \max(\underline{v}(\theta), \underline{v}(\varphi))$  by theorem 29  $= \max(v_*(\theta), v_*(\varphi))$  by induction  $v_*(\theta \vee \varphi) = v_*(\psi)$  by the properties of Tarski valuations. Also,  $\bar{v}(\psi) = \bar{v}(\theta \vee \varphi) = \max(\bar{v}(\theta), \bar{v}(\varphi))$  by the properties of supervaluation pairs as outlined above [31]  $= \max(v^*(\theta), v^*(\varphi))$  by induction  $= v^*(\theta \vee \varphi) = v^*(\psi)$  by the properties of Tarski valuations.

Part (ii): Proceed by induction on the complexity of sentences. For  $\psi = \neg p_i \in S\mathcal{L}^{-,0}$  then by definition 27 we have that  $\forall v \in \Pi, v(p_i) \leq v^*(p_i) \Rightarrow 1 - v(p_i) \geq 1 - v^*(p_i) \Rightarrow v(\neg p_i) \geq v^*(\neg p_i)$ . Hence,  $v^*(\neg p_i) = \min\{v(\neg p_i) : v \in \Pi\} = \underline{v}(\neg p_i)$ . Also,  $\forall v \in \Pi, \forall p_i \in \mathcal{P}, v(p_i) \geq v_*(p_i) \Rightarrow 1 - v(p_i) \leq 1 - v_*(p_i) \Rightarrow v(\neg p_i) \leq v_*(\neg p_i)$ . Hence,  $v_*(\neg p_i) = \max\{v(\neg p_i) : v \in \Pi\} = \bar{v}(\neg p_i)$ . The inductive steps then mirror those of part (i).  $\square$

If adopting the epistemic approach to uncertainty described in section 3 then a supervaluationist agent would represent her beliefs in terms of a probability distribution on  $\mathbb{V}_s$ . Similarly to Kleene belief pairs (definition 8) this naturally results in lower and upper belief measures on the sentences of  $\mathcal{L}$  of the form:  $\forall \theta \in S\mathcal{L}$ ;

$$\underline{\mu}(\theta) = w(\{\vec{v} \in \mathbb{V}_s : \underline{v}(\theta) = 1\}) \text{ and } \bar{\mu}(\theta) = w(\{\vec{v} \in \mathbb{V}_s : \bar{v}(\theta) = 1\})$$

In this case we refer to  $(\underline{\mu}, \bar{\mu})$  as a *supervaluation belief pair* or a *bounded supervaluation belief pair* if  $w$  is non-zero only on  $\mathbb{V}_{bs}$ . It is straightforward to show that the lower and upper measures comprising a supervaluation belief pair are respectively Dempster-Shafer belief and plausibility measures on  $S\mathcal{L}$  [44]. To see this notice that  $(\underline{\mu}, \bar{\mu})$  can be rewritten in Shafer's well-know mass function notation in the following manner. Let  $m : 2^{\mathbb{V}_c} \rightarrow [0, 1]$  be such that if  $\Pi \subseteq \mathbb{V}_c$  is the set of admissible valuations for  $\vec{v}$  then  $m(\Pi) = w(\vec{v})$ . Furthermore, if for  $\theta \in S\mathcal{L}$  we let  $\Pi(\theta) = \{v \in \mathbb{V}_c : v(\theta) = 1\}$  then it holds that:

$$\underline{\mu}(\theta) = \sum_{\Pi \subseteq \Pi(\theta)} m(\Pi) \text{ and } \bar{\mu}(\theta) = \sum_{\Pi \cap \Pi(\theta) \neq \emptyset} m(\Pi)$$

The idea of linking Dempster-Shafer theory to an underlying truth model dates back to Jaffray [21], who proposed a betting semantics for belief functions based on, in our notation, the lower supervaluation  $\underline{v}$ . Jaffray's associated Dutch book theorem is then a special case of Paris's general result [41], and the type of bets proposed are equivalent to lower bets [31], as described in section 5, but in a supervaluationist rather than a Kleene setting. The relationship between supervaluationism and Dempster-Shafer theory has also been noted by Field [15], who defines probability measures over the sentences of a modal logic language with an operator  $D$  denoting 'determinate'. In this setting the belief value of a sentence  $\theta$  corresponds to the probability of  $D\theta$ .

In contrast to Kleene belief pairs (theorem 9), supervaluation belief pairs are not in general additive, and instead the lower measure is super-additive and the upper measure is sub-additive. In [31], however, bounded supervaluation belief pairs are shown to be equivalent to Kleene belief pairs when restricted entirely to the sentences in  $S\mathcal{L}^+ \cup S\mathcal{L}^-$ . Furthermore, for sentences outside this class, the bounded supervaluation pairs are more precise than the corresponding Kleene belief pairs, in that they allocate lower probability values to borderline cases.

We can now adapt definition 10 in order to introduce truth degrees as corresponding to the mid-point of supervaluation belief pairs. For truth degrees of this kind generated from bounded supervaluation belief pairs, there is a close relationship with probability measures on  $S\mathcal{L}$  as shown in the following theorem.

**Theorem 32.** *Let  $w$  be a probability distribution on  $\mathbb{V}_{bs}$  then there exists a probability distribution  $w'$  on  $\mathbb{V}_c$  with an associated probability measure  $P$  on  $S\mathcal{L}$  given by  $\forall \theta \in S\mathcal{L}$ ,  $P(\theta) = w'(\{v \in \mathbb{V} : v(\theta) = 1\})$ , such that  $\forall \theta \in S\mathcal{L}^+ \cup S\mathcal{L}^-$ ,*

$$P(\theta) = td(\theta) = \frac{\underline{\mu}(\theta) + \overline{\mu}(\theta)}{2} \text{ where}$$

$$\underline{\mu}(\theta) = w(\{\vec{v} \in \mathbb{V}_{bs} : \underline{v}(\theta) = 1\}) \text{ and } \overline{\mu}(\theta) = w(\{\vec{v} \in \mathbb{V}_{bs} : \overline{v}(\theta) = 1\})$$

*Proof.* Let  $\{\vec{v} \in \mathbb{V}_{bs} : w(\vec{v}) > 0\} = \{\vec{v}_1, \dots, \vec{v}_m\}$  and then define  $w'$  on  $\mathbb{V}_c$  such that:  $\forall v \in \mathbb{V}_c$ ,

$$w'(v) = \sum_{i: v_{i*}=v} \frac{w(\vec{v}_i)}{2} + \sum_{i: v_i^*=v} \frac{w(\vec{v}_i)}{2}$$

Now for  $\theta \in S\mathcal{L}^+$  it holds that:

$$\begin{aligned} td(\theta) &= \frac{\underline{\mu}(\theta) + \overline{\mu}(\theta)}{2} = \frac{w(\{\vec{v}_i : \underline{v}_i(\theta) = 1\})}{2} + \frac{w(\{\vec{v}_i : \overline{v}_i(\theta) = 1\})}{2} \\ &= \frac{w(\{\vec{v}_i : v_{i*}(\theta) = 1\})}{2} + \frac{w(\{\vec{v}_i : v_i^*(\theta) = 1\})}{2} \text{ by theorem 31} \\ &= w'(\{v : v(\theta) = 1\}) = P(\theta) \end{aligned}$$

Similarly, for  $\theta \in S\mathcal{L}^-$  it holds that:

$$\begin{aligned} td(\theta) &= \frac{\underline{\mu}(\theta) + \overline{\mu}(\theta)}{2} = \frac{w(\{\vec{v}_i : \underline{v}_i(\theta) = 1\})}{2} + \frac{w(\{\vec{v}_i : \overline{v}_i(\theta) = 1\})}{2} \\ &= \frac{w(\{\vec{v}_i : v_i^*(\theta) = 1\})}{2} + \frac{w(\{\vec{v}_i : v_{i*}(\theta) = 1\})}{2} \text{ by theorem 31} \\ &= w'(\{v : v(\theta) = 1\}) = P(\theta) \end{aligned}$$

as required. □

Hence, theorem 32 shows that the truth degree generated from a bounded supervaluation belief pair corresponds to a probability measure on  $S\mathcal{L}^+ \cup S\mathcal{L}^-$ . Furthermore, we have the following corollary:

**Corollary 33.** Let  $w$  be a probability distribution on  $\mathbb{V}_{bs}$  such that  $\{\vec{v} : w(\vec{v}) > 0\} = \{\vec{v}_1, \dots, \vec{v}_m\}$  where  $\vec{v}_1 \preceq \dots \preceq \vec{v}_m$  then the credibility measure defined by

$$td(\theta) = \frac{\underline{\mu}(\theta) + \bar{\mu}(\theta)}{2}$$

satisfies the following:  $\forall \theta, \varphi \in S\mathcal{L}^+(S\mathcal{L}^-)$ ;

$$td(\theta \wedge \varphi) = \min(td(\theta), td(\varphi)) \text{ and } td(\theta \vee \varphi) = \max(td(\theta), td(\varphi))$$

*Proof.* Since  $\vec{v}_1 \preceq \dots \preceq \vec{v}_k$  then it follows that:

$$v_{1*} \trianglelefteq v_{2*} \trianglelefteq \dots \trianglelefteq v_{k*} \trianglelefteq v_k^* \trianglelefteq v_{k-1}^* \trianglelefteq \dots \trianglelefteq v_1^*$$

Define  $w'$  on  $\{v_{1*}, \dots, v_{k*}, v_k^*, \dots, v_1^*\}$  as in the proof of theorem 32 then the result follows immediately from theorem 25.  $\square$

**Example 34.** Consider a language  $\mathcal{L}$  with propositional variables  $\mathcal{P} = \{p_1, p_2, p_3, p_4, p_5\}$ . We now define the Tarski valuations  $v^{(i)} \in \mathbb{V}_c$  for  $i = 1, \dots, 6$  as given in the following table:

|           | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ |
|-----------|-------|-------|-------|-------|-------|
| $v^{(1)}$ | 0     | 0     | 0     | 0     | 0     |
| $v^{(2)}$ | 1     | 1     | 0     | 0     | 0     |
| $v^{(3)}$ | 0     | 0     | 1     | 1     | 0     |
| $v^{(4)}$ | 0     | 0     | 0     | 0     | 1     |
| $v^{(5)}$ | 0     | 0     | 1     | 1     | 1     |
| $v^{(6)}$ | 1     | 1     | 1     | 1     | 1     |

Based on these classical valuations, we now define supervaluation pairs  $\vec{v}_1, \dots, \vec{v}_4$  as characterised by the following sets of admissible valuations:

$$\begin{aligned} \Pi_1 &= \{v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)}, v^{(5)}, v^{(6)}\}, \quad \Pi_2 = \{v^{(1)}, v^{(3)}, v^{(4)}, v^{(5)}\}, \\ \Pi_3 &= \{v^{(1)}, v^{(4)}\}, \quad \Pi_4 = \{v^{(4)}\} \end{aligned}$$

In addition, we define a probability distribution on supervaluation pairs such that  $w(\vec{v}_1) = 0.5$ ,  $w(\vec{v}_2) = 0.3$ ,  $w(\vec{v}_3) = 0.1$  and  $w(\vec{v}_4) = 0.1$ . Now notice that  $\vec{v}_1, \dots, \vec{v}_4$  are bounded supervaluation pairs which have the following bounding classical valuations:

$$v_{1*} = v^{(1)}, v_{2*} = v^{(1)}, v_{3*} = v^{(1)}, v_{4*} = v^{(4)}, v_4^* = v^{(4)}, v_3^* = v^{(4)}, v_2^* = v^{(5)}, v_1^* = v^{(6)}$$

Hence, the conditions of theorem 32 are satisfied and we can define the following probability distribution  $w'$  on  $\mathbb{V}_c$ :

$$\begin{aligned} w'(v^{(1)}) &= \frac{w(\vec{v}_1)}{2} + \frac{w(\vec{v}_2)}{2} + \frac{w(\vec{v}_3)}{2} = 0.25 + 0.15 + 0.05 = 0.45 \\ w'(v^{(4)}) &= \frac{w(\vec{v}_4)}{2} + \frac{w(\vec{v}_4)}{2} + \frac{w(\vec{v}_3)}{2} = 0.05 + 0.05 + 0.05 = 0.15 \\ w'(v^{(5)}) &= \frac{w(\vec{v}_2)}{2} = 0.15, \quad w'(v^{(6)}) = \frac{w(\vec{v}_1)}{2} = 0.25 \end{aligned}$$



In this case for  $\neg p_2 \vee \neg p_3 \in S\mathcal{L}^-$ , we have that:

$$P(\neg p_2 \vee \neg p_3) = w'(v^{(1)}) + w'(v^{(4)}) + w'(v^{(5)}) = 0.45 + 0.15 + 0.15 = 0.75$$

Also,

$$\begin{aligned}\underline{\mu}(\neg p_2 \vee \neg p_3) &= w(\vec{v}_2) + w(\vec{v}_3) + w(\vec{v}_4) = 0.5 \text{ and} \\ \bar{\mu}(\neg p_2 \vee \neg p_3) &= w(\vec{v}_1) + w(\vec{v}_2) + w(\vec{v}_3) + w(\vec{v}_4) = 1 \text{ therefore} \\ td(\neg p_2 \vee \neg p_3) &= \frac{0.5 + 1}{2} = 0.75\end{aligned}$$

On the other hand for  $\neg p_1 \wedge p_5 \notin S\mathcal{L}^+ \cup S\mathcal{L}^-$ , we have that:

$$P(\neg p_1 \wedge p_5) = w'(v^{(4)}) + w'(v^{(5)}) = 0.15 + 0.15 = 0.3$$

But,

$$\begin{aligned}\underline{\mu}(\neg p_1 \wedge p_5) &= w(v^{(4)}) = 0.1 \text{ and } \bar{\mu}(\neg p_1 \wedge p_5) = w(\vec{v}_1) + w(\vec{v}_2) + w(\vec{v}_3) + w(\vec{v}_4) = 1 \\ \text{therefore } td(\neg p_1 \wedge p_5) &= \frac{0.1 + 1}{2} = 0.55\end{aligned}$$

In addition, notice that  $\vec{v}_1 \preceq \vec{v}_2 \preceq \vec{v}_3 \preceq \vec{v}_4$ , and hence the conditions of corollary 33 are satisfied so that, for example, we have the following:

$$\begin{aligned}\underline{\mu}(\neg p_2) &= 0.5, \bar{\mu}(\neg p_2) = 1 \Rightarrow td(\neg p_2) = 0.75 \\ \underline{\mu}(\neg p_3) &= 0.2, \bar{\mu}(\neg p_3) = 1 \Rightarrow td(\neg p_3) = 0.6\end{aligned}$$

Hence,

$$\max(td(\neg p_2), td(\neg p_3)) = \max(0.75, 0.6) = 0.75 = td(\neg p_2 \vee \neg p_3)$$

In general, supervaluation belief pairs in which  $w$  is non-zero only on a sequence of increasingly sharp valuations, fall within the scope of possibility theory [9]. More specifically, in such cases the lower and upper beliefs are, respectively, necessity and possibility measures on  $S\mathcal{L}$ . This is clear if we adopt the mass function notation outlined above, since by theorem 30 the mass will now only be non-zero on a nested sequence of sets of classical valuations  $\Pi_m \subseteq \Pi_{m-1} \subseteq \dots \subseteq \Pi_1$ , where  $\Pi_i$  is the set of admissible valuations for  $\vec{v}_i$ . In this case truth degrees are *credibility measures* as first proposed by Dubois and Prade [9] and later developed at some length by Liu and Liu [36]. The additional assumption of boundedness required in theorem 32 and corollary 33 does not seem to be typical in possibility theory.

## 7 Conclusions

In this paper we have identified clear links between probability theory in a three-value truth setting and min-max fuzzy logic. The most complete bridge is formed when the underlying

truth model is Kleene’s three-valued logic. In this case fully compositional fuzzy truth degrees as defined on a finite propositional language, can be completely characterised in terms of probability distributions of Kleene valuation pairs. In a classical (Tarski) or a supervaluationist context, the min-max calculus for truth degrees can be recaptured only on the language fragments  $SL^+$  and  $SL^-$ . Across all of these truth models the min-max operators result when there is a natural semantic precision ordering on the truth valuations, and in particular when the only uncertainty is about the relative vagueness or sharpness of the interpretation of  $\mathcal{L}$ . The latter, however, is a strong assumption and is likely not to hold in many of the contexts in which an agent must allocate subjective belief to sentences involving vague propositions. Indeed, we would go so far as to claim that if truth degrees are interpreted probabilistically then truth-functionality will only arise as a result of strong assumptions which are in turn only ever likely to be applicable in relatively restricted contexts. In order to model more complex scenarios than those consistent with a fully compositional calculus, there is then a case for the integrated study of vagueness and uncertainty within a richer and less restrictive representational framework. This would be a new direction distinct from the formal study of fuzzy logics as truth-functional systems [19] and which would have a quite different motivation.

A number of authors have highlighted the somewhat confusing relationship between fuzziness and vagueness. For example, Zadeh [57] is adamant that they are completely distinct phenomena. Dubois [13] emphasises the differences between gradualness and vagueness, where a gradual predicate induces a partial ordering on the underlying conceptual domain, whilst vagueness is epitomised by explicit borderline cases. In contrast, we adopt a broader more encompassing view of vagueness as a multifaceted phenomenon and where vague predicates may exhibit any, or all, of the three main symptoms identified by Keefe [22]. From this perspective our work has identified a clear bridge between fuzziness and two aspects of vagueness; namely explicit borderlines and blurry (uncertain) boundaries. In a nutshell, our proposed semantics identifies fuzziness with subjective probabilities of vague sentences, so that truth degrees are determined by underlying probabilities of the three truth values **t**, **f** and **b** as illustrated in figure 3. Furthermore, the valuation pair framework has provided us with a common notation with which to explore probability in both a Kleene and a supervaluationist setting; these being two widely proposed approaches to truth-gaps in the literature on vagueness. As already mentioned above, in both theories the relationship to min-max fuzzy logic is defined in terms of the semantic precision (or sharpening) ordering. This ordering relates directly to the explicit borderline aspect of vagueness, with one valuation being vaguer than another if it permits more borderline cases.

In general, vagueness is frowned upon in science and engineering, where clarity and semantic precision are seen as being a fundamental prerequisite to progress. A hypothesis must be precisely formulated before it can be properly empirically tested. From this

perspective, if subjective probability is seen as being a normative theory of belief for ideally rational agents, then there is no justification for considering its application to vague propositions. On the other hand, given its ubiquitousness in natural language one is strongly inclined to suspect that vagueness must have a positive role to play in complex multi-agent communication systems. To quote Lipman [34], "It seems rather far fetched to conclude that we have simply tolerated a worldwide, several thousand year efficiency loss". Recently, Van Deemter [50] has proposed a number of communication tasks in which vagueness can be useful. These include *search*, where typicality information embedded in vague predicates can be exploited in order to reduce search times [51], and *future contingencies*, referring to the use of vagueness to mitigate the risk of making promises or forecasts. Indeed, [31] has proposed a decision model based on Kleene belief pairs which directly exploits borderline cases so as to reduce assertion risks. In Lawry and Dubois [32] it is proposed that vagueness has a role to play in allowing agents to reach a consensus between them, whilst at the same time maintaining some level of internal consistency within their own beliefs. Furthermore, [32] also proposed several consensus combination operators for Kleene belief pairs in this context. Another positive role for vagueness, at least in the guise of semantic uncertainty, is in language learning. In the introduction we have already outlined a case for the explicit representation of semantic uncertainty given the empirical nature of language learning. In this context, O' Connor [38] shows that under time limited conditions and for high dimensional state spaces, optimal learning methods incorporate vagueness. Also, in agent based systems, [14] has shown how populations of agents playing a simple language game can evolve a shared set of vague categories which are then effective in communications. Overall then there is an emerging, although admittedly still rather embryonic, case for the utility of vagueness in communications. From this pragmatic viewpoint it can be perfectly rational for an agent to use vague predicates as a descriptive and representational tool.

## Acknowledgements

The authors would like to thank the anonymous referees for their insightful comments and suggestions.

## References

- [1] N.D. Belnap, (1977), 'A Useful Four-Valued Logic', in *Modern Uses of Multiple Valued Logics* (Eds: J.M. Dunn and G. Epstein), Reidel, Dordrecht
- [2] P. Cheeseman, (1985), 'In Defense of Probability' *Proceedings of IJCAI 85*, Vol. 2, pp1002-1009.

- [3] P. Cheeseman, (1986), ‘Probabilistic versus fuzzy reasoning’ in (Eds: L. Kanal, J. Lemmer), *Uncertainty in Artificial Intelligence*, North-Holland, pp85-102.
- [4] D. Ciucci, (2011), ‘Orthopairs: A Simple and Widely Used Way to Model Uncertainty’, *Fundamenta Informaticae*, Vol. 108, pp287-304.
- [5] D. Ciucci, D. Dubois, J. Lawry, (2012), ‘Borderline vs. Unknown: a Comparison Between Three-Valued Valuations, Partial Models, and Possibility Distributions’, *Working Papers from the ECAI Workshop on Weighted Logics for Artificial Intelligence*, pp83-90.
- [6] R. Cooke, (2004), ‘The Anatomy of the Squizzel: The Role of Operational Definitions in Representing Uncertainty’, *Reliability Engineering and Systems Safety*, Vol. 85, pp313-319.
- [7] B. de Finetti, (1931), ‘Sul significato soggettivo della probabilità’, *Fundamenta Mathematicae*, Vol. 17, pp298-329.
- [8] D. Dubois, H. Prade, (1982), ‘On Several Representations of an Uncertain Body of Evidence’, in (Eds: M.M. Gupta and E. Sanchez) *Fuzzy Information and Decision Processes*, North-Holland, pp167-181
- [9] D. Dubois, H. Prade, (1988), *Possibility Theory*, Plenum Press.
- [10] D. Dubois, H. Prade, (1993), ‘Fuzzy Sets and Probability : Misunderstandings, Bridges and Gaps’, *Second IEEE Conference of Fuzzy Systems*, pp1059-1068.
- [11] D. Dubois, H. Prade, (1997), ‘The Three Semantics of Fuzzy Sets’, *Fuzzy Sets and Systems*, Vol. 90, pp141-150.
- [12] D. Dubois, (2008), ‘On Ignorance and Contradiction Considered as Truth-Values’, *Logic Journal of the IGPL*, Vol. 16, No. 2, pp195-216.
- [13] D. Dubois, (2011), ‘Have Fuzzy Sets Anything to do with Vagueness?’, in (Eds. P. Cintula, C. G. Fermüller, L. Godo, P. Hájek) *Understanding Vagueness: Logical Philosophical and Linguistic Perspectives*, Studies in Logic, Vol. 36, College Publications, pp311-333.
- [14] H. Eyre, J. Lawry, (2013), ‘Communication success in vague category games’, *Proceedings of AISB 2013*.
- [15] H. Field, (2000), ‘Indeterminacy, Degree of Belief, and Excluded Middle’, *Nous*, Vol. 34, No. 1, pp1-30.
- [16] K. Fine, (1975), ‘Vagueness, Truth and Logic’, *Synthese*, Vol. 30, pp265-300.

- [17] R. Giles, (1988), ‘The Concept of Grade of Membership’, *Fuzzy Sets and Systems*, Vol. 25, pp297-323.
- [18] I.R. Goodman, (1984), ‘Some New Results Concerning Random Sets and Fuzzy Sets’, *Information Science*, Vol. 34, pp93-113
- [19] P. Hájek, (1998), *Metamathematics of Fuzzy Logic*, Trends in Logic Vol. 4, Springer.
- [20] E. Hisdal, (1988), ‘Are grades of membership probabilities’, *Fuzzy Sets and Systems*, Vol. 25, pp325-348.
- [21] J-Y. Jaffray, (1989), ‘Coherent Bets Under Partially Resolving Uncertainty and Belief Functions’, *Theory and Decision*, Vol. 26, pp99-105.
- [22] R. Keefe, P. Smith (Eds), (2002), *Vagueness: A Reader*, MIT Press.
- [23] S.C. Kleene, (1952), *Introduction to Metamathematics*, D. Van Nostrand Company Inc., Princeton, New Jersey.
- [24] G. J. Klir, B. Yuan, (1995), *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall.
- [25] D. Lassiter, (2011), ‘Vagueness as Probabilistic Linguistic Knowledge’, *Lecture Notes in Artificial Intelligence*, Vol. 6517, pp127-150.
- [26] M. Laviolette, J. W. Seaman Jr, (1994), ‘The Efficacy of Fuzzy Representations of Uncertainty’, *IEEE Transactions on Fuzzy Systems*, Vol.2, No.1, pp4-15.
- [27] J. Lawry, (2008), ‘Appropriateness Measures: An Uncertainty Model for Vague Concepts’, *Synthese*, Vol. 161(2), pp255-269.
- [28] J. Lawry, Y. Tang, (2009), ‘Uncertainty modelling for vague concepts: A prototype theory approach’, *Artificial Intelligence*, Vol. 173, pp1539-1558.
- [29] J. Lawry, I. Gonzalez-Rodriguez, (2011), ‘A Bipolar Model of Assertability and Belief’, *International Journal of Approximate Reasoning*, Vol. 52, pp76-91.
- [30] J. Lawry, (2010), ‘Imprecise Bipolar Belief Measures Based on Partial Knowledge from Agent Dialogues’, *Lecture Notes in Computer Science*, Vol. 6379, pp205-218.
- [31] J. Lawry, Y. Tang, (2012), ‘On Truth-gaps, Bipolar Belief and the Assertability of Vague Propositions’, *Artificial Intelligence*, Vol. 191-192, pp20-41.
- [32] J. Lawry, D. Dubois, (2012), ‘A Bipolar Framework for Combining Beliefs about Vague Propositions’, *Proceedings of 13th International Conference on Principles of Knowledge Representation and Reasoning*, pp530-540.

- [33] J. Lawry, T.P. Martin, (2013), ‘Conditional Beliefs in a Bipolar Framework’, *Lecture Notes in Computer Science*, Vol. 7958, pp364-375.
- [34] B. L. Lipman, (2006), ‘Why is language vague?’, *Working Paper*, Department of Economics, Boston University.
- [35] D. V. Lindley, (1987), ‘The Probability Approach to the Treatment of Uncertainty in Artificial Intelligence and Expert Systems’, *Statistical Science*, Vol. 12, No. 1, pp17-24.
- [36] B. Liu, Y-K. Liu, (2002), ‘Expected Value of Fuzzy Variable and Fuzzy Expected Value Models’, *IEEE Transactions on Fuzzy Systems*, Vol. 10, No. 4, pp445-450.
- [37] V. J. Loginov, (1966), ‘Probability Treatment of Zadeh Membership Functions and Their Use in Pattern Recognition’, *Engineering Cybernetics*, pp68-69.
- [38] C. O’Connor, (2013), ‘The Evolution of Vagueness’, *Erkenntnis*, in press, DOI 10.1007/s10670-013-9463-2.
- [39] R. Parikh, (1996), ‘Vague Predicates and Language Games’, *Theoria(Spain)*, Vol XI(27), pp97-107.
- [40] J.B. Paris, (1994), *The Uncertain Reasoners Companion: A Mathematical Perspective*, Cambridge University Press.
- [41] J. B. Paris, (2001), ‘A Note on the Dutch Book Method’, *Proceedings of the Second International Symposium on Imprecise Probabilities and their Applications*, Ithaca, New York.
- [42] G. Priest, (2001), *An Introduction to Non-classical Logics*, Cambridge University Press.
- [43] D. Ripley, (2011), ‘Contradictions at the Borders’, *Lecture Notes in Artificial Intelligence*, Vol. 6517, pp169-188.
- [44] G. Shafer, (1976), *A Mathematical Theory of Evidence*, Princeton University Press.
- [45] S. Shapiro, (2006), *Vagueness in Context*, Oxford University Press.
- [46] N. D. Singpurwalla, J. M. Booker, (2004), ‘Membership Functions and Probability Measures of Fuzzy Sets’, *Journal of the American Statistical Association*, Vol. 99, pp867-877.
- [47] S. Soames, (1998), *Understanding Truth*, Oxford University Press.

- [48] R. Scozzafava, (2013), ‘The Membership of a Fuzzy Set as Coherent Conditional Probability’, in (Eds. R. Seising, E. Trillas, C. Moraga, S. Termini) *Fuzziness*, Studies in Fuzziness and Soft Computing, Vol. 299, Springer, pp631-635.
- [49] M. Tye, (1994), ‘Sorites Paradoxes and the Semantics of Vagueness’, *Philosophical Perspectives*, Vol. 8, pp189-206.
- [50] K. van Deemter, (2009), ‘Utility and Language Generation: The Case of Vagueness’, *Journal of Philosophical Logic*, Vol. 38, pp607-632.
- [51] K. van Deemter, (2010), ‘Vagueness Facilitates Search’, in (Eds. M. Aloni, H. Bastiaanse, T. de Jager, K. Schulz) *Logic, Language and Meaning*, Lecture Notes in Computer Science, Vol. 6042, pp173-182.
- [52] P. Walley, (1991), *Statistical Inference from Imprecise Probabilities*, Chapman and Hall, London.
- [53] J. Robert G. Williams, (2013), ‘Probability and Non-classical Logic’, forthcoming in (Eds. Hajek and Hitchcock), *Oxford Companion to the Philosophy of Probability*, OUP.
- [54] T. Williamson, (1994), *Vagueness*, Routledge.
- [55] L. A. Zadeh,(1965), ‘Fuzzy Sets’, *Information and Control*, Vol.8, No.3 pp338-353.
- [56] L. A. Zadeh, (1975), ‘The Concept of a Linguistic Variable and its Application to Approximate Reasoning:I’, *Information Sciences*, Vol. 8, pp199-249.
- [57] L. A. Zadeh, (1978), ‘PRUF-A Meaning Representation Language for Natural Languages’, *International Journal of Man-Machine Studies*, Vol. 10, pp395-460.